

**PCT**ORGANISATION MONDIALE DE LA PROPRIÉTÉ INTELLECTUELLE  
Bureau international

## DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIERE DE BREVETS (PCT)

(51) Classification internationale des brevets <sup>6</sup> : <b>C12N 15/48, C12Q 1/70, C07K 14/15, A61K 31/70</b>		<b>A1</b>	(11) Numéro de publication internationale: <b>WO 99/67395</b> (43) Date de publication internationale: 29 décembre 1999 (29.12.99)
(21) Numéro de la demande internationale: <b>PCT/FR99/01513</b> (22) Date de dépôt international: <b>23 juin 1999 (23.06.99)</b> (30) Données relatives à la priorité: <b>98/07920 23 juin 1998 (23.06.98) FR</b> (71) Déposant (pour tous les Etats désignés sauf US): <b>INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE-INSERM [FR/FR]; 101, rue de Tolbiac, F-75654 Paris Cedex 13 (FR).</b> (72) Inventeurs; et (75) Inventeurs/Déposants (US seulement): <b>ALLIEL, Patrick, M. [FR/FR]; 4, rue Lazare Carnot, F-92140 Clamart (FR). PERIN, Jean-Pierre [FR/FR]; 182, rue d'Aulnay, F-92350 Le Plessis-Robinson (FR). RIEGER, François [FR/FR]; 38 bis, boulevard de la République, F-92100 Boulogne (FR).</b> (74) Mandataire: <b>CABINET ORES; 6, avenue de Messine, F-75008 Paris (FR).</b>		(81) Etats désignés: <b>AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, brevet ARIPO (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), brevet eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</b>  <b>Publiée</b> <i>Avec rapport de recherche internationale.</i> <i>Avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues.</i>	
(54) Title: <b>NUCLEIC SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HUMAN ENDOGENOUS RETROVIRAL MOTIFS, AND THEIR USES</b>			
(54) Titre: <b>FAMILLE DE SEQUENCES NUCLEIQUES ET DE SEQUENCES PROTEIQUES DEDUITES PRESENTANT DES MOTIFS RETROVIRAUX ENDOGENES HUMAINS, ET LEURS APPLICATIONS</b>			
(57) Abstract <p>The invention concerns a novel nucleic sequence and deduced protein sequence family with whole or partial human endogenous retroviral motifs. The invention also concerns the detection and/or the use of said nucleic sequences and said corresponding protein sequences or fragments of said sequences, for diagnostic, prophylactic and therapeutic uses, in particular for neuropathological conditions with autoimmune constituent such as multiple sclerosis. Said purified nucleic acid sequences comprise all or part of a sequence coding for a human endogenous retroviral sequence having at least <i>env</i>-type retroviral motifs, corresponding to the sequence SEQ ID NO:1 or to a sequence having a homology level with said sequence SEQ ID NO:1 not less than 80 % on more than 190 nucleotides or not less than 70 % on more than 600 nucleotides for <i>env</i>-type domains. The invention further concerns the use of the flanking or adjacent sequences of said sequences and controlled by the latter, as diagnostic reagents.</p>			
(57) Abrégé <p>Nouvelle famille de séquences nucléiques et de séquences protéiques déduites, qui présentent des motifs rétroviraux endogènes humains complets ou partiels. Détection et/ou utilisation desdites séquences nucléiques et desdites séquences protéiques correspondantes ou de fragments de ces séquences, dans le cadre d'applications diagnostiques, prophylactiques et thérapeutiques, en particulier pour des neuropathologies à composante auto-immune comme la sclérose en plaques. Lesdites séquences d'acide nucléique purifiées comprennent tout ou partie d'une séquence codant pour une séquence rétrovirale endogène humaine, qui présente au moins des motifs rétroviraux de type <i>env</i>, répondant à la séquence SEQ ID NO:1 ou à une séquence présentant un niveau d'homologie avec ladite séquence SEQ ID NO:1 supérieur ou égal à 80 % sur plus de 190 nucléotides ou supérieur ou égal à 70 % sur plus de 600 nucléotides pour les domaines de type <i>env</i>. Utilisation des séquences flanquantes ou adjacentes desdites séquences et contrôlées par ces dernières, comme réactifs de diagnostic.</p>			

# *UNIQUEMENT A TITRE D'INFORMATION*

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïdjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TC	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave de Macédoine	TM	Turkménistan
BF	Burkina Faso	GR	Grèce	ML	Mali	TR	Turquie
BG	Bulgarie	HU	Hongrie	MN	Mongolie	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MR	Mauritanie	UA	Ukraine
BR	Brésil	IL	Israël	MW	Malawi	UG	Ouganda
BY	Bélarus	IS	Islande	MX	Mexique	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	NE	Niger	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NL	Pays-Bas	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norvège	YU	Yougoslavie
CH	Suisse	KG	Kirghizistan	NZ	Nouvelle-Zélande	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire démocratique de Corée	PL	Pologne		
CM	Cameroun	KR	République de Corée	PT	Portugal		
CN	Chine	KZ	Kazakhstan	RO	Roumanie		
CU	Cuba	LC	Sainte-Lucie	RU	Fédération de Russie		
CZ	République tchèque	LI	Liechtenstein	SD	Soudan		
DE	Allemagne	LK	Sri Lanka	SE	Suède		
DK	Danemark	LR	Libéria	SG	Singapour		
EE	Estonie						

## FAMILLE DE SEQUENCES NUCLEIQUES ET DE SEQUENCES PROTEIQUES DEDUITES PRESENTANT DES MOTIFS RETROVIRAUX ENDOGENES HUMAINS, ET LEURS APPLICATIONS

5                   La présente invention est relative à une nouvelle famille de séquences nucléiques et de séquences protéiques déduites, qui présentent des motifs rétroviraux endogènes humains complets ou partiels, ainsi que des séquences flanquantes ou adjacentes desdites séquences, et contrôlées par ces dernières : modification de l'expression ou altération de la structure (polyadénylation, épissage alternatif...) desdites séquences flanquantes.

10                   L'invention est également relative à la détection et/ou à l'utilisation desdites séquences nucléiques et desdites séquences protéiques correspondantes, dans le cadre d'applications diagnostiques, prophylactiques et thérapeutiques, en particulier pour des neuropathologies à composante autoimmune comme la sclérose en plaques.

15                   L'invention concerne aussi l'obtention de sondes nucléiques double brins et simple brin anti-sens, de ribozymes, aptes à moduler la répllication virale (T.R. Cech, *Science*, 1987, **236**, 1532-1539 ; R.H. Symons, *Trends Biochem. Sci.*, 1989, **14**, 445-450) des molécules recombinantes correspondantes, et des anticorps associés.

20                   Les rétrovirus sont des virus qui se répliquent uniquement en utilisant la voie inverse du traitement classique de l'information génétique. Ce processus, nommé transcription inverse, est médié par une ADN polymérase ARN dépendante ou transcriptase reverse, codée par le gène *pol*. Les rétrovirus codent aussi au minimum pour deux gènes additionnels. Le gène *gag* code pour les protéines du squelette, matrice, nucléocapside et capside. Le gène *env* code pour les glycoprotéines d'enveloppe.

25                   La transcription rétrovirale est régulée par des régions promotrices ou "enhancers", situées dans des régions hautement répétées ou LTR (*Long Terminal Repeat*) et qui sont présentes aux deux extrémités du génome rétroviral.

30                   Lors de l'infection d'une cellule, la polymérase fait une copie ADN du génome ARN ; cette copie peut alors s'intégrer dans le génome humain. Les rétrovirus ne tuent pas les cellules qu'ils infectent, mais au contraire améliorent souvent leur rapidité de croissance. Les rétrovirus peuvent infecter des cellules germinales ou

des embryons à un stade précoce ; ils peuvent dans ces conditions, intégrer la lignée germinale et être transmis par transmission mendélienne verticale. ce qui constitue la relation la plus étroite entre un hôte et son parasite. Ces virus endogènes peuvent dégénérer au cours des générations de l'organisme hôte et perdre leurs propriétés initiales. Cependant certains d'entre eux peuvent conserver tout ou partie de leurs propriétés ou des propriétés des motifs les composant, ou encore acquérir de nouvelles propriétés fonctionnelles présentant un avantage pour l'organisme hôte. ce qui expliquerait la préservation de leur séquence.

L'existence de motifs endogènes présentant de longs cadres de lecture ouverts et/ou soumis à une forte pression de sélection peut donc être indicatrice d'une fonction biologique préservée ou acquise, qui peut correspondre à un bénéfice pour l'organisme hôte. Ces séquences rétrovirales peuvent aussi subir, au cours des générations, des modifications discrètes qui vont être à même de réveiller certaines de leurs potentialités et engendrer ou favoriser des processus pathologiques. Il est apparu récemment nécessaire de faire le bilan et d'identifier ces séquences afin de pouvoir évaluer leur impact fonctionnel.

Les séquences rétrovirales endogènes humaines ou HERVs représentent une part importante du génome humain. Ces régions rétrovirales se présentent sous plusieurs formes :

- des structures rétrovirales endogènes complètes associant des motifs *gag*, *pol* et *env*, flanqués de séquences nucléiques répétées, qui montrent une analogie significative avec la structure LTR-*gag-pol-env*-LTR des rétrovirus infectieux,
- des séquences rétrovirales tronquées ; par exemple, les rétrotransposons sont privés de leur domaine *env* et les rétroposons ne possèdent pas les régions *env* et LTR.

Jusqu'à présent l'étude de ces régions du génome a été négligée chez l'Homme pour deux raisons essentielles :

- l'existence d'insertions/délétions qui peuvent décaler le cadre de lecture et de mutations qui modifient la séquence. Ces modifications entraînent des altérations de la structure et par conséquent de la fonction biologique de ces motifs.



- l'absence d'associations avérées avec des pathologies humaines.

La connaissance, récente de fragments significativement représentatifs du génome humain et une orientation des recherches vers une étude structure/fonction des motifs rétroviraux endogènes, ont permis de préciser l'intérêt de ces régions. L'implication de séquences endogènes tronquées ou complètes dans des pathologies chez l'animal est documentée ; par exemple leur association avec des processus tumoraux a été clairement mise en évidence (S.K. Chattopadhyay et coll., 1982, *Nature*, **295**, 25-31). Une recherche visant à préciser l'association ou l'influence des HERVs dans des pathologies humaines se justifie donc aujourd'hui.

10 Une classification des éléments HERV a été proposée (Tönjes R.R. et al., *AIDS & Hum. Retrovirol.*, 1996, **13**, S261-S267; A.M. Krieg et al., *FASEB J.*, 1992, **6**, 2537-2544). Elle est basée sur une homologie de ces séquences avec des rétrovirus isolés chez les animaux, à l'aide de sondes rétrovirales hétérologues. En effet, en général, les HERVs présentent relativement peu d'homologie avec des rétrovirus infectieux humains connus.

Les familles de classe I présentent une homologie de séquence avec les rétrovirus de mammifères de type C ; on peut citer notamment la superfamille ERI, proche du virus MuLV (*murine leukemia virus*) et du virus BaEV (*baboon endogenous virus*).

20 Les familles de classe II présentent une homologie de séquence avec les rétrovirus de mammifères de type B tel que le MMTV (*mouse mammary tumour virus*) ou les rétrovirus de type D tel que le SRV (*squirrel monkey retrovirus*).

D'autres familles ont également été décrites ; parmi celles-ci, on peut citer des HERVs qui présentent, de manière exceptionnelle, une homologie partielle avec HTLV-1 (RTVL-H) ou des virus de primates ; HRES-1, par exemple, présente une homologie de séquence avec des HTLVs.

25 Les programmes de très grand séquençage du génome humain permettent aujourd'hui de disposer d'un nombre significatif de nouvelles séquences rétrovirales. L'usage de logiciels de traitement de données permet d'identifier et d'analyser ces gènes. Dans ce contexte une recherche systématique portant sur l'ensemble des informations disponibles à ce jour a été engagée afin d'identifier de nouvelles séquences

rétrovirales endogènes humaines en fonction de certains critères d'analyse :

- présence de longs cadres de lecture ouverts conservés au cours de l'évolution de l'organisme hôte et pouvant laisser envisager une fonction biologique.

- analogie avec des séquences déjà caractérisées en dehors ou dans le  
5 domaine des rétrovirus,

- localisation dans des régions de susceptibilité pour certaines pathologies ou à proximité de gènes essentiels, par exemple dans les domaines du cancer, de la régulation du système immunitaire ou dans certaines neuropathologies.

Les recherches effectuées par les Inventeurs, dans des bases de données de séquences leur ont permis d'identifier un ensemble de séquences ou de motifs  
10 rétroviraux endogènes dont l'expression normale ou pathologique peut favoriser ou perturber un effet protecteur vis-à-vis de processus pathologiques, ou intervenir dans le déclenchement ou l'aggravation de pathologies.

La présente invention a pour objet un fragment d'acide nucléique  
15 purifié, caractérisé en ce qu'il comprend tout ou partie d'une séquence codant pour une séquence rétrovirale endogène humaine, qui présente au moins des motifs rétroviraux de type *env*, répondant à la séquence SEQ ID NO:1 ou à une séquence présentant un niveau d'homologie avec ladite séquence SEQ ID NO:1 supérieur ou égal à 80% sur plus de 190 nucléotides ou supérieur ou égal à 70 % sur plus de 600 nucléotides pour  
20 les domaines de type *env*.

On entend par séquence homologue, aussi bien une séquence qui présente une identité complète ou partielle avec la séquence SEQ ID NO:1 précitée qu'une séquence qui présente une similarité partielle avec ladite séquence SEQ ID NO:1.

25 Selon un mode de réalisation avantageux dudit fragment, il présente à la fois des motifs rétroviraux correspondant à un domaine *env* et répondant à la séquence SEQ ID NO:1 et des motifs rétroviraux correspondant à un domaine *gag* et répondant à la séquence SEQ ID NO:2 ou à une séquence présentant un niveau d'homologie supérieur ou égal à 80 % sur plus de 190 nucléotides ou supérieur ou  
30 égal à 70 % sur plus de 600 nucléotides pour les domaines de type *env* et un niveau d'homologie supérieur ou égal à 90 % sur plus de 700 nucléotides ou supérieur ou

égal à 70 % sur plus de 1200 nucléotides pour les domaines de type *gag*, lesquels motifs ne présentent aucune insertion ou délétion de plus de 200 nucléotides.

Lesdits fragments constituent une nouvelle famille de séquences rétrovirales endogènes humaines (famille HERV-7q) qui présente une homologie de  
5 séquence avec les rétrovirus MSRV, tels que décrits dans la Demande Internationale WO 97/06260 ; lesdits fragments selon la présente invention présentent :

- deux motifs nucléotidiques répétés de 711 pb (figure 3), présentant des signaux caractéristiques identifiés dans des LTRs (*Long Terminal Repeats*) : promoteurs de transcription de type boîtes TATAA ou CCAAT. Ces domaines répétés  
10 encadrent trois motifs déduits de type-*gag*, *pol* et *env* (figure 2).

- un motif de type *env* (positions 6965 nt - 9550 nt sur la séquence SEQ ID NO :3 ou sur la figure 1) qui contient un long cadre de lecture ouvert de 1620 nucléotides (positions 7874-9493 de la séquence ID NO:3 et figure 1), codant pour une protéine de séquence inédite de 540 acides aminés appelée envérine (figure 4 et  
15 SEQ ID NO:26) et fragment souligné de la figure 18. On retrouve à l'intérieur du domaine trans-membranaire de ce domaine *env*, un motif peptidique de type CKS-25/CKS-17 (figure 5), reconnu pour présenter des fonctions immunosuppressives sur les cellules lymphocytaires hôtes (M. Mitani et coll., 1987, *Proc. Natl. Acad. Sci. USA*, 84, 237-240). Un domaine de type doigt de zinc (*zinc-finger*)  $HX_3-HX_{22}$ .  
20  $_{33}CX_2C$  (Kulkolski et coll., 1992, *Mol. Cell. Biol.*, 12, 2331-2338), que l'on retrouve dans des domaines de type intégrase est identifié dans un autre cadre de lecture. Ce domaine *env* particulier signe la caractéristique de nouveaux motifs rétroviraux endogènes.

- le motif (positions 3065 nt - 4390 nt sur la séquence SEQ ID NO:3)  
25 de type-*gag* codant pour des motifs protéiques selon la figure 6 (SEQ ID NO:58) (positions 3118-4198 de la SEQ ID NO:3) a été identifié grâce à des analogies avec des domaines *gag* connus. On retrouve, par exemple, la région d'homologie majeure  $QX_3EX_7R$  (Benit et coll., 1997, *J. Virol.*, 71, 5652-5657). Le motif de fixation des acides nucléiques  $CX_2CX_3-HX_4C$ , situé en position C-terminale, est identifié dans un  
30 autre cadre de lecture (Covey et coll., 1986, *Nucleic Acids Res.*, 14, 623-633). En amont du domaine *gag* on détecte un motif de 182 nucléotides répété deux fois (figure

1).

- le domaine *pol* présente les consensus classiques d'une région *pol* de rétrovirus au niveau des domaines protéase, transcriptase reverse et RNase H. On retrouve dans *pol* un motif proche du consensus LLDTGA (Weber et coll., 1988, Science, 243, 928-931). Les motifs D et AF, LPQ et SP, et YVDD (Xiong et Eickbush, 1990, EMBO J., 9, 3353-3362), sont respectivement retrouvés dans les 3°, 4° et 5° boîtes d'homologie. Les motifs YTDGSS et TDS sont présents dans la région de la RNase H,

- les régions *gag* et *pol* pourraient être considérées comme jointives avec un passage de la région *gag* à la région *pol* par un décalage du cadre de lecture.

La présente invention englobe les séquences appartenant à la famille HERV-7q telle que définie ci-dessus (présence de la séquence SEQ ID NO:1 ou d'une séquence homologue ou présence à la fois des séquences SEQ ID NO:1 et SEQ ID NO:2) et notamment les séquences SEQ ID NO:3-22, 28 et 61 ; elle englobe également les séquences nucléiques complémentaires et les séquences inverses complémentaires des séquences précédentes ainsi que les fragments issus des régions codantes des séquences précédentes correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires. (SEQ ID NO :37-57, 59-60 et 121-122).

Ces différents fragments peuvent avantageusement être utilisés comme amorces ou comme sondes (réactifs A) ; ils s'hybrident spécifiquement dans des conditions de forte stringence à une séquence de la famille HERV-7q.

Parmi ces fragments, on peut citer, de préférence les fragments suivants:

- un fragment de 182 nucléotides répété deux fois, situé en amont du domaine *gag* aux positions 2502-2611/2613-2865 de la SEQ ID NO:3 ;

Amorces et sondes spécifiques de la région *gag*

- une amorce G1F, sens, localisée dans la région amont du domaine *gag* de HERV-7q : 5' GGACCATAGAGGACACTCCAGGACTA 3' (SEQ ID NO:37);

- une amorce G1R, anti-sens, localisée dans la région 3' terminale du

domaine *gag* : 5' CCTCAGTCCTGCTGCTGGATCATCT 3' (SEQ ID NO:38)

- le fragment de 1505 nt amplifié par le couple G1F-G1R est utilisé afin de générer les sondes aptes à hybrider les différents produits d'amplification des PCR ;

- 5                   - une amorce G2F, sens nichée : (SEQ ID NO:39)  
                   5' CCTCCAAGCAGTGGGAGGAAGAGAATT 3'  
                   - une amorce G2R, anti-sens nichée : (SEQ ID NO:40)  
                   5' CCTTCCCTGTGTTATTGTGGACATCATT 3'  
                   - une amorce G4F, sens nichée : (SEQ ID NO:41)  
 10               5' GGAAGAAGTCTATGAATTATTCAATGATGT 3'  
                   - une amorce G3F, sens nichée: (SEQ ID NO:42)  
                   5' GGGACACAGAATCAGAACATGGAGATT 3'  
                   - une amorce G4R, anti-sens nichée : (SEQ ID NO:43)  
                   5' GCCTTCAGAAGAGTCAGGTGACAGAGA 3'  
 15               - une amorce G5R, anti-sens nichée : (SEQ ID NO:44)  
                   5' GAGCCTCCAAAGTCCACTTGCCTGA 3'

Amorces et sondes spécifiques de la région *env*

- une amorce E1F, sens : (SEQ ID NO:45)  
                   5' GATTTCAGTATCTACTAGTCTGGGTAGAT 3'  
 20               - une amorce E1R, anti-sens : (SEQ ID NO:46)  
                   5' CTAGGAAATCCAGCTAGTCCTGTCTCA 3'  
                   - le fragment de 2529 nt amplifié par le couple d'amorces E1F-E1R,  
 est utilisé afin de générer les sondes aptes à hybrider les différents produits  
 d'amplification des PCR.

- 25               - une amorce E2F, sens : (SEQ ID NO:47)  
                   5' CCAAGACAGCCAACTTAGTTGCAGACAT 3'  
                   - une amorce E2R, antisens : (SEQ ID NO:48)  
                   5' GGACGCTGCATTCTCCATAGAAACTCTT 3'  
                   - une amorce E3F, sens : (SEQ ID NO:49)  
 30               5' GCAATACTACATACACAACCAACTCCCAA 3'

- une amorce E3R, anti-sens : (SEQ ID NO:50)  
5' GGGGGAGGCATATCCAACAGTTAGTA 3'
- une amorce E4F, sens : (SEQ ID NO:51)  
5' CCATCTACACTGAACAAGATTTATACACTT 3'
- 5 - une amorce E4R, anti-sens : (SEQ ID NO:52)  
5' AATGCCAGTACCTAGTGCACCTAGCACT 3'
- une amorce E5F, sens : (SEQ ID NO:53)  
5' CGAATACAACGTAGAGCAGAGGAGCTTCGAA 3'
- une amorce E6F, sens : (SEQ ID NO:54)  
10 5' AGCCCAAGATGCAGTCCAAGACTAAGAT 3'
- une amorce E5R : (SEQ ID NO:55)  
5' GCGTAGTAGAGGTTGTGCAGCTGAGAT 3'
- une amorce ExF : (SEQ ID NO:56)  
CCCTTACCAAGAGTTTCTATGGAGAAT
- 15 - une amorce ExR : (SEQ ID NO:57)  
ACCGCTCTAACTGCTTCCTGCTGAATT

Tous les oligonucléotides sont conçus pour pouvoir générer une amorce sens et une amorce anti-sens par un décalage de la séquence de l'amorce de référence de 1 à 7 nucléotides vers le côté 5' ou vers le côté 3': la modification de la

20 séquence peut entraîner une modification de la taille de l'amorce de 1 à 7 nucléotides selon les cas. Les amorces choisies peuvent être optimisées selon les cas par un raccourcissement ou un allongement portant sur 1 à 9 nucléotides.

De manière préférée, l'hybridation, le clonage, le sous-clonage, l'obtention, la préparation et l'analyse des acides nucléiques, des peptides et des anti-

25 corps, le séquençage des acides nucléiques et des peptides, l'hybridation *in situ* et l'immunohistochimie sont réalisés dans les conditions décrites dans les ouvrages suivants :

- Current Protocols in Molecular Biology. Eds. F.M Ausubel, R. Brent & R.E Kingston et coll. Green Publishing associates and Wiley Interscience.
- 30 - Molecular Cloning: a laboratory manual. Eds. J. Sambrook, E.F. Fritsch & T. Maniatis. Cold Spring Harbor Laboratory Press. Cold Spring Harbor.

- The Practical Approach series. Eds. D. Rickwood & B.D. Ames. IRL Press and Oxford University Press. En particulier, antibodies I & II; DNA cloning I, II, III; Nucleic acid and protein sequence analysis; Nucleic acid hybridization; Nucleic acid sequencing ; Oligonucleotide synthesis; Protein purification applications;
- 5 Protein purification methods; Protein sequencing; Transcription and translation: Gels electrophoresis of nucleic acids; Gels electrophoresis of proteins; Genome analysis; HPLC of macromolecules; Human genetic diseases; Microcomputing in biology; Molecular neurobiology; Mutagenicity testing; Essential molecular biology I & II.
- Proteome research: New frontiers in functional genomics. Eds
- 10 M.R. Wilkins & coll.. Springer.

La séquence rétrovirale endogène humaine (SEQ ID NO:3). située sur le bras long du chromosome 7 correspond à la séquence HERV-7q ; elle présente 10,5 kb (fig. 1 et 2) et répond aux critères précédemment définis.

- La recherche de domaines présentant des similitudes, tout ou partie,
- 15 avec les régions *gag* et *env* de HERV-7q a abouti à l'identification de nouvelles séquences rétrovirales endogènes. Ces séquences peuvent présenter la structure d'un rétrovirus endogène complet comme la séquence rétrovirale endogène située à proximité du gène des sous-unités alpha et delta du récepteur des cellules-T, et dénommée en conséquence HERV-TcR ; à titre d'exemple la figure 7 montre la comparaison des
- 20 alignements nucléiques des domaines *gag* respectifs de HERV-7q et HERV-TcR (séquence HG12, SEQ ID NO:19). On trouve aussi des structures rétrovirales partielles. Ces domaines rétroviraux similaires à HERV-7q sont identifiées dans des séquences nucléiques indépendantes comme le montre leur localisation chromosomique. Des motifs nucléiques (appelés ici, HEx ou HGx et respectivement analogues à des domaines de type *env* ou *gag*) ressemblant aux domaines *env* ou *gag* de
- 25 HERV-7q ont été retrouvés, à l'aide des banques de données précitées :

- HE2 : chromosome 17 (SEQ ID NO:4),
- HE3 et HG3: chromosome 6 (SEQ ID NO:5 et 6),
- HE4 : chromosome X (SEQ ID NO:7),
- 30 - HE5 : chromosome X q22 (SEQ ID NO:8),
- HE6 et HG6 : chromosome 1 q23.3-q24.3 (SEQ ID NO:9 et 10),

- HE7 : chromosome 7 p15 (SEQ ID NO:11),
- HE8 et HG8 : chromosome 19 (SEQ ID NO:12 et 13),
- HE9 : chromosome X (SEQ ID NO:14),
- HE10 : chromosome X q13.1-21.1 (SEQ ID NO:15),
- 5 - HE11 et HG11 : chromosome 7 q21-22 (SEQ ID NO:16 et 17),
- HE12 et HG12, dans HERV-TcR : chromosome 14 q11.2 (SEQ ID NO:18 et 19),
- HE13 (SEQ ID NO:61) : chromosome 6 q24.1-24.3

La présente invention englobe également les fragments codants et  
 10 non codants pour tout ou partie de l'envérine comprenant au moins 14 nucléotides et  
 notamment les fragments codant pour la partie C-terminale de l'envérine, soit à partir  
 de l'acide aminé 291, soit à partir de l'acide aminé 321, à compter de la première  
 méthionine.

Ces fragments comprennent en particulier une zone critique où deux  
 15 insertions de 12 nucléotides ont été caractérisées :

- une première insertion a été identifiée (séquence A), chez des indi-  
 vidus de 2 groupes (malades et témoins). Cette insertion située entre les acides aminé  
 487 et 488, permet d'insérer le térapeptide VLQM. Une analyse comparative montre  
 que cette insertion est identifiée dans une région homologue située dans la séquence  
 20 HE13, appartenant à la famille HERV-7q. L'amplification de la séquence de type  
 HE13, pourrait indiquer qu'il existe une altération de la séquence de l'envérine de  
 HERV-7q, ce qui favoriserait l'amplification de la séquence contenue dans HE13.  
 Cette observation permet aussi d'utiliser cette insertion comme élément spécifique  
 d'amplification de séquences de type HE13.

25 Une deuxième insertion (séquence B) a été identifiée chez un patient  
 présentant une SEP. L'insertion de 12 nucléotides est située au niveau de l'acide  
 aminé 495 et code pour le térapeptide MQSM. Il est remarquable de constater que  
 cette insertion est aussi identifiée dans une région homologue située dans HE13.

Séquence A: TAAACTACAAATGGTTCTTCAAATGGAGCCCA  
 30 (SEQ ID NO:59)



Séquence B: GATGCAGTCCAAGATGTCAGTCCATGACTAAGA  
(SEQ ID NO:60).

Ces observations mettent en évidence des modifications de la séquence de l'envérine de type HERV-7q qui constituent la base d'une stratégie de  
5 détection par amplification spécifique d'allèles (AS-PCR), permettant de détecter ces différences dans une population et qui pourraient correspondre, soit à une mutation/délétion associée à une certaine susceptibilité, soit à un polymorphisme, soit à une mutation/délétion associée à une pathologie comme la sclérose en plaques.

Les alignements des domaines *env* (fig. 8) et *gag* (fig. 9) explicitent  
10 les niveaux d'homologie observés entre les séquences décrites ci-dessus et les séquences homologues dans HERV-7q. Les analogies peuvent s'étendre aux motifs rétroviraux flanquants.

Une analyse des séquences étiquettes disponibles dans les banques de données montre que des transcrits appartenant à certains des membres de cette  
15 famille, en particulier HERV-7q, s'expriment essentiellement dans des tissus d'origine fœtale ou placentaire.

Des séquences polypeptidiques générées par ces transcrits peuvent donc être potentiellement produites et des fonctions ou activités biologiques peuvent être envisagées, par analogie avec des polypeptides biologiquement actifs d'origine  
20 virale ou rétrovirale ; par exemple, les motifs peptidiques de type CKS-17 (Haraguchi et al., PNAS, 1995, 92, 5568-5571) (fig. 5) ou CKS-25 (Huang S.S et Huang J.S, J. Biol. Chem. 1998, 273, 4815-4818), qui présentent des fonctions immunomodulatrices sur les cellules lymphocytaires hôtes. Les différences de séquence observées et d'éventuelles modifications normales ou pathologiques, sont en particulier, à l'origine  
25 d'une modulation de la fonction.

HERV-7q représente le paradigme de la nouvelle famille de séquences rétrovirales endogènes humaines ou de motifs rétroviraux endogènes.

HERV-7q et certaines des séquences rétrovirales endogènes appartenant à sa famille, présentent un domaine de type *pol* analogue à des séquences rétrovirales de type *pol* comme par exemple la région *pol* identifiée dans le rétrovirus  
30 MSRV associé à la sclérose en plaques et décrit par H. Perron et al. (1997, *Proc. Natl.*

*Acad. Sci. USA*, 94, 7583-7588 ; Demande Internationale PCT WO 97/06260).

Toutefois, les séquences selon la présente invention se distinguent des séquences rétrovirales exogènes infectieuses analogues à MSRV antérieurement décrites en ce que les séquences *gag* et *env*, selon l'invention sont significativement  
5 différentes selon les critères précédemment définis et en fonction de certaines caractéristiques spécifiques, par exemple le long cadre de lecture ouvert du domaine *env* de HERV-7q ; elles seraient à même de permettre de signer une pathologie lorsqu'elles présentent des insertions, des délétions, des décalages de cadre de lecture ou des mutations.

10 En effet, les différences observées entre les séquences humaines de type HERV-7q, qui sont isolées d'individus réputés normaux et les séquences issues de certains échantillons d'origine pathologique, ne sont pas distribuées au hasard. Des comparaisons menées entre la région *gag* provenant de particules rétrovirales infectieuses (N° d'accension EMBL: A60168, A60200, A60201, A60171...) et la séquence  
15 *gag* correspondante de HERV-7q (fig. 9), permettent d'observer que les mutations affectent préférentiellement des codons non-sens. Par exemple, deux codons non-sens dans HERV-7q sont remplacés par un codon arginine dans A60200, ce qui permet d'obtenir une séquence déduite de 109 acides aminés pour HERV-7q et de 166 acides aminés pour A60200. Les changements de base permettent en conséquence de prolonger le cadre de lecture et de coder potentiellement pour des structures polypeptidiques  
20 de plus grande taille (figure 10).

De même, une séquence de type *env* provenant de particules rétrovirales infectieuses, présente une analogie significative avec le domaine *env* de HERV-7q (figure 11). Ces analogies marquées entre séquences rétrovirales exogènes  
25 et endogènes pourraient être à l'origine du déclenchement ou de l'aggravation de certains processus pathologiques, en particulier de certaines maladies auto-immunes, comme la sclérose en plaques. A cet égard, on peut remarquer que certaines des séquences rétrovirales endogènes décrites dans l'invention se situent à proximité ou dans des régions réputées présenter une susceptibilité pour la sclérose en plaques : par  
30 exemple HERV-7q et la région 7q21-22 du chromosome 7, de même pour HE12 et HG12 dans HERV-TcR et la région du gène codant pour les chaînes alpha et delta du

récepteur des cellules-T, HE2 et le chromosome 17, ou HE3, HE13 et HG3 et le chromosome 6, par exemple, les séquences HE11 et HG11, autour de la région 7q 21-22 ou encore HE4, HE5, HE6, HE9, HE10 ou HG10 sur le chromosome X. Ces séquences seraient donc à même de fournir des moyens de localisation ou  
5 d'identification des gènes de prédisposition.

On n'observe aucune homologie significative avec des séquences rétrovirales endogènes déjà décrites, par contre, on peut relever une homologie limitée, permettant d'identifier une structure générale de domaine *env* : toutefois, ladite homologie est inférieure aux critères définis selon l'invention entre les  
10 domaines *env* de la séquence HERV-7q (SEQ ID NO :1) et de la séquence HERV-9 (figure 12). La figure 11 montre des homologies étendues entre la séquence HERV-7q avec une séquence rétrovirale exogène (N° d'accension EMBL : A60170).

Les séquences rétrovirales endogènes humaines appartenant à la famille de HERV-7q, peuvent protéger contre des agressions liées à l'environnement  
15 ou constituer un bénéfice pour l'individu. Cet effet bénéfique pourrait être une des raisons possibles de la pression de sélection exercée sur certaines de ces séquences et du caractère potentiellement fonctionnel des structures protéiques déduites identifiées : par exemple le long cadre de lecture ouvert apte à coder pour une nouvelle protéine et correspondant au domaine *env* de HERV-7q.

20 Les séquences rétrovirales endogènes humaines appartenant à la famille de HERV-7q pourraient être associées par exemple, à des pathologies en relation avec les processus liés au cancer, aux neuropathologies à composante auto-immune ou à tout autre processus pathologique en association ou non avec des virus ou rétrovirus endogènes ou exogènes. Leur action pourrait porter sur la déclaration, l'aggravation, la  
25 modification du calendrier d'apparition ou encore la protection vis à vis de la maladie.

Dans le contexte d'application à des pathologies autoimmunes (comme par exemple le lupus, le syndrome de Sjögren, la polyarthrite rhumatoïde, la sclérose en plaques...), on peut relever des analogies significatives entre les motifs rétroviraux endogènes identifiés et des motifs retrouvés dans des structures rétrovirales  
30 caractérisées chez des patients présentant des pathologies autoimmunes comme la sclérose en plaques : par exemple des fragments de domaine *gag* (récemment dispo-

nibles dans les banques de données) provenant de particules rétrovirales infectieuses ou encore la séquence complète du domaine *pol* correspondant au virus MSRV associé à la sclérose en plaques. Ces motifs rétroviraux possèdent des analogies significatives avec les séquences endogènes homologues de type HERV-7q, ce qui permet  
5 d'envisager une association directe ou indirecte avec des processus pathologiques, dont la sclérose en plaques, en association ou non avec MSRV.

L'intérêt de ces séquences dépasse le cadre des maladies auto-immunes. En dehors de l'importance générale des motifs rétroviraux dans le déclenchement ou l'aggravation d'un processus tumoral, bien montré en particulier dans les  
10 modèles murins (H. Fan dans *The retroviridae*, 1994, ed. J.A. Levy, Plenum. New York, p. 313-353), ces séquences pourraient se retrouver à proximité ou au sein de gènes importants et en altérer l'expression : par exemple HERV-TcR et les gènes des sous-unités alpha et delta du récepteur des cellules T impliquées dans des perturbations de la fonction immunitaire.

15 La présente invention englobe, en outre, l'utilisation de séquences associées aux séquences de la famille HERV-7q pour la détection et/ou le pronostic de différentes maladies auto-immunes (neuropathologies, en particulier) ; ces séquences codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou l'altération (polyadénylation, épissage alternatif) est associée à l'expression normale  
20 ou pathologique ou à la régulation/dérégulation des motifs appartenant à la famille HERV-7q et correspondent à des transcrits ou des ADNc des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes ou encadrant des séquences rétrovirales de la famille HERV-7q.

On entend par région flanquante, toute région située à proximité  
25 (incluse dans ou incluant) une séquence rétrovirale endogène appartenant à la famille HERV-7q, telle que définie ci-dessus, jusque et y compris les gènes immédiatement contigus et/ou situés à une distance ne pouvant excéder 120 kb.

Les Inventeurs ont maintenant trouvé que la présence des séquences rétrovirales telles que définies ci-dessus, perturbent l'expression ou altèrent la struc-  
30 ture des séquences flanquantes définies ci-après.

Les transcrits desdites séquences flanquantes (et leurs fragments,

notamment ceux soulignés ou en italique dans les figures 14-16, 22-26, sont définies ci-après :

- à 1021 pb en amont de HERV-7q, on identifie une séquence rétrovirale endogène appelée RH7 (SEQ ID NO:62 et figure 22) ; cette séquence est située en 5' de la séquence HERV-7q ; dans la figure 22, la partie en italique correspond au début de la séquence HERV-7q ; la séquence RH7 est soulignée : deux sites de polyadénylation putatifs sont en gras. Cette séquence SEQ ID NO:62 présente une homologie significative, sur plus de 6 kb, avec des séquences rétrovirales endogènes de type RGH (figure 13). Des séquences appartenant à cette famille s'expriment en particulier chez des patients présentant une arthrose rhumatoïde (Nakagawa et coll., (1997), Arthritis. Rheum., 40, 627- 638). La présente invention inclut également des fragments de la séquence SEQ ID NO:62, comprenant entre 14 et 50 nucléotides (utilisation comme amorces), de préférence entre 14 et 25 nucléotides ou au moins 25 nucléotides (utilisation comme sonde), lesquels fragments présentent les caractéristiques suivantes : les 4 nucléotides de l'extrémité 3' sont différents des motifs correspondant de la séquence RGH2 (séquence du bas dans la figure 13, n° accession à GenBank : D110 18).

- à moins de 9 kb en amont de HERV-7q, on identifie la séquence RAM75 (SEQ ID NO:63 et figure 14) contenant les 24 exons codants (qui couvrent près de 41 kb), du gène de l'ATPase péroxysomale PEX1. PEX1, en association avec PEX6 est responsable de l'importation des protéines péroxysomales et de la stabilisation du récepteur PEX5. Une perturbation/altération affectant PEX1 est responsable de diverses neuropathologies comme le syndrome de Zellweger, l'adrénoleucodystrophie néonatale et la forme infantile de la maladie de Refsum (Reuber et coll., (1997), Nature Genet., 17, 445- 448). On peut rappeler que la fonction principale des péroxysomes est associée au métabolisme des acides gras, en particulier par des processus de  $\beta$ -oxydation. Une altération du gène identifié dans la séquence RAM75 ou de son expression, par modification de la fonction des régions 5' et 3' régulatrices ou encore par modification des épissages ou des processus de polyadénylation, en particulier sous l'influence de motifs rétroviraux voisins, seraient à même de perturber l'expression ou la structure de l'ATPase et par conséquent de perturber une des

fonctions péroxysoniales, en particulier le métabolisme des lipides, en particulier myéliniques, avec des conséquences pour certaines pathologies, dont des neuropathologies, comme la sclérose en plaques ; les parties soulignées (figure 14) correspondent aux 24 exons codants.

5 La présente invention inclut également les fragments de la séquence SEQ ID NO:63, inclus dans les 24 exons codants précités et comprenant au moins 14 nucléotides.

L'analyse du profil d'expression (transcrits et protéines) de la séquence RAM75 (SEQ ID NO:63) est un bon indicateur du diagnostic différentiel des  
10 neuropathologies à composante auto-immune.

A la figure 14, les exons codants sont soulignés. Les codons d'initiation et non-sens ainsi que les sites putatifs de polyadénylation sont en gras et soulignés.

- à 0.7 kb en aval de la séquence HERV-7q et sur près de 17 kb  
15 (SEQ ID NO:64 et figure 15), on identifie la séquence nucléotidique RAV73, où l'on détecte des séquences étiquettes et des exons potentiels aptes à produire une ou plusieurs séquences polypeptidiques ; l'invention inclut également des fragments de cette séquence SEQ ID NO:64 compris dans les séquences étiquettes et les exons potentiels tels qu'ils apparaissent (parties soulignées) à la figure 15, lesquels  
20 fragments comportent au moins 14 nucléotides.

- à 120 kb en amont de la séquence HG3, et sur 15 kb, se trouve la séquence nucléotidique RBP3 (SEQ ID NO:65 et figure 23), qui couvre l'extrémité 3'du gène codant pour un facteur de transcription de la famille Blimp-1 (SEQ ID NO:119 et figure 25), une protéine de 789 acides aminés qui est un répresseur de  
25 l'expression du gène de l'interféron-béta (Keller et Maniatis, Genes Dev., (1991), 5, 868-879), qui est déjà associé à certaines pathologies malignes (Mock et coll., Genomics, (1996), 37, 24-28), et qui pourrait jouer un rôle dans la différenciation et la pathogenèse des cellules B. L'intérêt de l'association possible de la séquence rétrovirale endogène contenant les motifs HG3 et HE3 et de Blimp-1 est multiple, dans le  
30 cas de pathologies, et en particulier la sclérose en plaques. Blimp-1 agit en particulier sur les cellules B dont on connaît la contribution dans les processus inflammatoires

associés à la sclérose en plaques. Blimp-1 est capable de bloquer l'induction virale du promoteur  $INF\beta$  dont on connaît l'aptitude à réduire la fréquence des poussées et la progression lésionnelle chez des patients atteints de SEP. Une perturbation de l'expression ou de la structure de Blimp-1, en relation avec un élément rétroviral de type HERV-7q, est associée en conséquence à des neuropathologies ou à des maladies à caractère auto-immun, comme la sclérose en plaques ; cette séquence nucléotidique RBP3 (SEQ ID NO:65) contient des motifs nucléotidiques identifiés dans la séquence nucléaire codant pour le gène Blimp-1 ; l'invention inclut aussi la détection des séquences ARNm de la protéine Blimp-1 (SEQ ID NO:119).

10 - la séquence rétrovirale endogène de type HERV-7q, contenant HE3 et HG3, se trouve située dans la région HI3 correspondant à un intron s'étendant sur plus de 46 kb (SEQ ID NO:66), d'un gène qui pourrait coder pour l'analogue d'APS (figure 24), une protéine de 275 acides aminés spécifique d'apoptose, surexprimée dans différents cellules en culture après déclenchement d'un processus apoptotique  
15 (Hammond et coll., FEBS Lett., (1998), 425, 391- 395). L'intron se situe au niveau de l'acide aminé 231 d'APS. L'extrémité de HE3 est à plus de 12 kb de l'extrémité 5' de l'intron, alors que HG3 est situé à plus de 28 kb de l'extrémité 3' de l'intron. Des processus apoptotiques sont associés à la sclérose en plaques. En particulier, il a été décrit un processus apoptotique affectant des astrocytes et des oligodendrocytes en  
20 présence d'une fraction purifiée de liquide céphalo-rachidien de patients atteints de sclérose en plaques (Ménard et coll., J. Neurol. Sci., (1998), 154, 209- 221).

Enfin, il faut souligner que la région nucléaire contenant HE3, HG3, HI3 et RBP3 est localisée au niveau du bras court du chromosome 6, en 6p21, qui est une région proposée de susceptibilité à la sclérose en plaques (The Multiple Sclerosis  
25 Genetic Group, Nature Genet., (1996), 13, 469- 472).

L'interaction entre les séquences de type HERV-7q et les séquences flanquantes et l'importance de l'établissement d'un profil d'expression incluant une ou plusieurs des séquences précitées pour établir un diagnostic différentiel d'une neuropathologie apparaît encore plus du fait que l'on observe que les séquences HG12 et  
30 HE12 sont situées dans une région intronique du gène codant pour les sous-unités alpha et delta des récepteurs des cellules T. Les récepteurs des cellules T sont impli-

qués dans les processus de régulation immunitaire et leur influence a été proposée dans le cas de maladies auto-immunes, dont la sclérose en plaques.

L'invention a également pour objet les transcrits générés à partir des séquences précitées ainsi que celles présentant éventuellement des modifications avec  
5 les séquences de référence décrites dans l'invention lorsqu'ils sont exprimés chez certains patients.

En effet, les systèmes de régulation de l'expression des protéines rétrovirales de HERV-7q, qui sont présents dans les motifs de type LTR, pourraient influencer l'expression de gènes situés dans le voisinage chromosomique proche ou  
10 éloigné et induire des perturbations à caractère immunologique et/ou neurologique. Par exemple la séquence rétrovirale endogène HERV-TcR, se trouve à proximité immédiate des gènes des sous-unités alpha et delta du récepteur des cellules-T précédemment décrit. Les motifs de type LTR pourraient aussi coder pour des superantigènes (Acha-Orbea et Palmer, 1991, *Immunol. Today*, 12, 356-361). D'une manière  
15 générale des protéines rétrovirales de type HERV-7q ou apparenté, ou leurs formes tronquées ou partielles pourraient être impliquées dans des phénomènes de cytotoxicité ou de superantégenicité, comme par exemple celles issues du long cadre de lecture ouvert identifié dans le domaine *env* (figure 4).

Des séquences du type des LTR 5' et 3' de HERV-7q, fortement  
20 conservées sont concernées par de tels effets régulateurs. A titre d'exemple on décrit LTX, une séquence comparable à celle d'une LTR de HERV-7q (SEQ ID NO:67 et figure 16), et qui se trouve au cœur d'un intron de plus de 49 kb, mais à 2 kb du site 5' donneur, du gène de FMR2 associé au X-fragile et codant pour une protéine de 1311 acides aminés (figure 26). Les LTR modulent l'épissage alternatif (Kapitonov et  
25 Jurka, (1999), *J. Mol. Evol.*, 48, 248- 251), l'expression du gène, la fixation sur des protéines nucléaires (Akopov et coll., (1998), *FEBS Lett.*, 421, 229- 233), ou permettent l'obtention d'un signal de polyadénylation alternatif (Goodchild et coll., (1992) , *Gene*, 121, 287- 294).

D'une manière générale, on peut remarquer l'existence de plusieurs  
30 séquences rétrovirales endogènes de type HERV-7q (HE4, HE5, HE9, HE10), situées au niveau du chromosome X qui représente le chromosome associé au plus grand



nombre de pathologies.

A cet égard, on peut relever que des motifs rétroviraux issus de régions défectives sont aptes à présenter des fonctions biologiques: par exemple, la protéine d'enveloppe p15E issue de motifs rétroviraux défectifs, possède une activité  
5 anti-inflammatoire et immunosuppressive (Snyderman et Ciancolo, 1984, *Immunol. Today*, 5, 240-244).

Ces structures sont vraisemblablement à même de provoquer des brèches ou d'amplifier des dérégulations dans les processus de défense immunitaire. Certains des motifs des domaines *gag*, *env* et de type LTR peuvent être associés à une  
10 fonction particulière ou peuvent contribuer à la fonction normale ou pathologique des domaines flanquants tels que définis ci-dessus (SEQ ID NO:62-67). Des recombinaisons avec un élément d'origine exogène, rétroviral ou non, peut donner lieu à la production de motifs nucléiques ou protéiques qui pourraient soit protéger, soit déclencher, ou favoriser ou aggraver une pathologie. De même, une structure rétro-  
15 virale contenant des éléments rétroviraux endogènes selon l'invention seraient à même de provoquer un processus pathologique après passage par un cycle transitoire exogène puis réintégration dans une région sensible ou critique du génome humain.

On peut ainsi obtenir des profils d'expression (transcrits et éventuellement protéines) qui correspondent aux neuropathologies précitées.

20 De même, la combinaison de motifs appartenant à la famille de HERV-7q, ou d'éléments induits par des motifs appartenant à la famille de HERV-7q, avec des motifs d'origine ou induits de manière exogène seraient à même de pouvoir déclencher, ou aggraver un processus pathologique ou au contraire de favoriser une protection ou une rémission partielle ou une guérison totale et définitive.

25 La détection rendue possible des domaines de type HERV-7q, suggère des applications possibles à la fois au niveau prophylactique, du pronostic et du diagnostic: par exemple des approches immunologiques ou d'amplification génique permettant de comparer des individus normaux servant de référence avec des patients, seraient à même de favoriser le dépistage, d'améliorer la détection précoce de  
30 la déclaration de la maladie et/ou de suivre l'évolution d'une pathologie chez des patients pouvant présenter une susceptibilité ou ayant déclaré la maladie ou encore

chez des individus considérés comme normaux, selon les critères cliniques actuels.

Les sondes nucléiques et immunologiques spécifiques, telles que définies, dans la présente invention sont à même de favoriser l'identification et la détection de motifs anormalement exprimés dans le cadre de pathologies associées au cancer, ou de neuropathologies en particulier auto-immunes, au premier rang  
5 desquelles la sclérose en plaques.

La présente invention a également pour objet des séquences nucléiques hybrides, caractérisées en ce qu'elles comprennent des séquences ou motifs appartenant à la famille de HERV-7q, ou d'éléments induits par des motifs appartenant à la famille de HERV-7q, avec des motifs d'origine ou induits de manière  
10 exogène (séquences rétrovirales exogènes) ; de telles séquences hybrides sont vraisemblablement à même de pouvoir déclencher, ou aggraver un processus pathologique ou au contraire de favoriser une protection ou une rémission partielle ou une guérison totale et définitive.

La présente invention a également pour objet un réactif de diagnostic pour la détection différentielle de séquences nucléiques endogènes humaines complètes ou partielles, présentant des motifs rétroviraux, sélectionnés parmi les séquences SEQ ID NO :1 et/ou SEQ ID NO :2, caractérisé en ce qu'il est sélectionné dans le groupe constitué par les séquences SEQ ID NO:1-22, 28, 37-57, 59-61 et 121-  
20 122, les séquences nucléiques complémentaires et les séquences inverses complémentaires des séquences précédentes, par les fragments nucléotidiques capables de définir ou d'identifier les séquences SEQ ID NO:1 et/ou SEQ ID NO:2 et toute séquence flanquante ou les chevauchant ainsi que par les fragments issus des régions codantes des séquences SEQ ID NO:1-22 et 61, correspondant à un cadre glissant  
25 supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires, éventuellement marquées avec un marqueur approprié ainsi que par les séquences telles que définies aux figures 18-21.

Les séquences des sondes nucléiques, ribonucléiques et oligonucléotidiques utilisées seront choisies dans les régions *env* et *gag* ou leur régions  
30 flanquantes : par exemple les oligonucléotides amorces pour HERV-7q, seront choisis dans les régions situées entre les nucléotides 3065 et 4390, les nucléotides 6965 et

9550 ou les nucléotides 2502-2865 de la SEQ ID NO:3, ainsi que dans toute séquence adjacente (amont ou aval) capable de permettre une amplification spécifique (figure 1).

Parmi les marqueurs appropriés, on peut citer, les isotopes radio-actifs, les enzymes, les fluorochromes, des marqueurs chimiques (biotine), les haptènes (digoxygénine) et les anticorps ou analogues de bases appropriées.

De manière préférée :

- ledit réactif est sélectionné parmi les séquences SEQ ID NO:37-57 et est apte à être utilisé comme amorce.

10 - ledit réactif est sélectionné parmi les séquences suivantes :

un fragment de 1505 nt amplifié par le couple d'amorces SEQ ID NO:37 et SEQ ID NO:38 (amorces G1F et G1R),

un fragment de 2529 nt amplifié par le couple d'amorces SEQ ID NO:45 et SEQ ID NO:46 (amorces E1F et E1R),

15 un fragment de 182 nucléotides répété deux fois, situé en amont du domaine *gag* aux positions 2502-2611/2613-2865,

des fragments codants ou non-codants pour tout ou partie de l'envérine, comprenant au moins 14 nucléotides et notamment les fragments codant pour la partie C-terminale de l'envérine, soit à partir de l'acide aminé 291, soit à partir  
20 de l'acide aminé 321, à compter de la première méthionine,

et est apte à être utilisé comme sonde.

La présente invention a également pour objet un procédé de détection rapide et différentiel des séquences nucléiques rétrovirales endogènes de type *env* ou *env* et *gag*, de leurs variants normaux ou pathologiques, par hybridation et/ou  
25 amplification génique, réalisé à partir d'un échantillon biologique, lequel procédé est caractérisé en ce qu'il comprend :

(a) une étape dans laquelle l'on met en contact un échantillon biologique à analyser avec au moins une sonde telle que définie ci-dessus et

(b) une étape dans laquelle on détecte par tout moyen approprié, le  
30 ou les produits résultant de l'interaction séquence nucléotidique-sonde.

Conformément audit procédé, il peut comprendre :

\* préalablement à l'étape (a) :

. une étape de préparation du tissu ou du liquide biologique concerné,

. une étape d'extraction de l'acide nucléique à détecter, et

5 . au moins un cycle d'amplification génique et

\* postérieurement à l'étape (b) :

. une étape de comparaison des séquences nucléiques obtenues dans ledit échantillon biologique avec les séquences rétrovirales endogènes humaines selon l'invention par tout moyen approprié et notamment par séquençage, Southern-blot, 10 coupure de restriction, SSCP ou toute autre méthode permettant d'identifier une insertion ou une délétion ou encore une simple mutation entre les différentes séquences comparées.

Conformément à l'invention, les séquences rétrovirales endogènes humaines selon l'invention sont ainsi comparées aux séquences nucléiques présentes 15 dans l'échantillon biologique à analyser et permettent la détection de séquences homologues de patients atteints de pathologies, susceptibles de mettre en jeu une modification de leur génome.

De manière avantageuse, lesdites comparaisons géniques sont menées à partir d'ADN génomique provenant d'individus témoins et de patients.

20 Une amplification génique classique par PCR sera menée à l'aide d'amorces 5' -sens et 3' -antisens encadrant ou comprenant la zone à étudier (zone *env* ou zone *gag*).

Également de manière avantageuse, les séquences des sondes nucléiques, ribonucléiques et oligonucléotidiques utilisées sont choisies dans les 25 régions *env* et *gag* ou leurs régions flanquantes : par exemple les oligonucléotides amorces pour HERV-7q, seront choisis dans les régions situées entre les nucléotides 3065 et 4390 et les nucléotides 6965 et 9550, ainsi que dans toute séquence adjacente (amont ou aval) capable de permettre une amplification spécifique (figure 1), comme précisé ci-dessus. Elles sont de préférence sélectionnées dans le groupe constitué par 30 un fragment de 1505 nt amplifié par le couple d'amorces SEQ ID NO:37 et SEQ ID NO:38 (amorces G1F et G1R),

un fragment de 2529 nt amplifié par le couple d'amorces SEQ ID NO:45 et SEQ ID NO:46 (amorces E1F et E1R).

L'étape d'amplification génique est notamment réalisée à l'aide d'une des techniques d'amplification génique suivante : amplification par la Q $\beta$ -réplicase, PCR, LCR, ERA, CPR ou SDA.

La présente invention a également pour objet des séquences chimères, caractérisées en ce qu'elles sont constituées par un fragment de 17 à 40 nucléotides d'une séquence flanquante telle que définie ci-dessus associée à un motif rétroviral endogène de type HERV-7q comprenant entre 17 et 40 nucléotides, tel que défini ci-dessus.

La présente invention a également pour objet un procédé de détection des transcrits, tels que définis ci-dessus, caractérisé en ce qu'il comprend :

- le prélèvement des ARN messagers provenant d'échantillons biologiques (tissus, cellules, fluides biologiques) témoins et d'un échantillon analogue prélevé chez des patients et

- l'analyse qualitative et/ou quantitative desdits ARNm, par hybridation *in situ*, par dot-blot, Northern-blot, RNase mapping ou RT-PCR, à l'aide d'un réactif de diagnostic tel que défini ci-dessus.

La présente invention a également pour objet une méthode de détection et/ou d'évaluation d'une sur-expression/sous-expression ou d'une modification d'au moins l'une des séquences ou fragments de séquences rétrovirales endogènes de type HERV-7q et/ou de leurs séquences flanquantes associées, caractérisée en ce qu'elle comprend :

- le dépôt sur un support approprié comme par exemple un filtre de nylon, une lame de verre ou leur équivalent, de l'ADNc ou son équivalent provenant de clones, de produits de PCR obtenus à partir d'ADN génomique, de produits de RT-PCR provenant de transcrits ou encore de séquences oligonucléotidiques spécifiques, lesdites séquences d'ADN étant des séquences ou des fragments de séquences rétrovirales endogènes de type HERV-7q et/ou leurs séquences flanquantes, telles que définies ci-dessus, constituées par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation

ou l'altération est associée à l'expression normale ou pathologique ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences correspondant à des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite  
5 famille HERV-7q et dont l'une des extrémités ne peut se trouver à une distance excédant 120 kb, et/ou une séquence chimère telle que définie ci-dessus,

- l'hybridation dudit support avec au moins une sonde marquée de manière adéquate obtenue, par exemple, par rétrotransposition d'un mélange d'ARN provenant de cellules, de tissus ou de liquides biologiques provenant de témoins  
10 réputés normaux, de membres de populations ethniques différentes, de patients atteints de pathologies souvent associées à une expression de rétrovirus, comme les processus tumoraux, ou comme les maladies auto-immunes, et

- la détection des hybrides formés.

Selon un mode de mise en œuvre avantageux de ladite méthode,  
15 ledit transcrit ou ADNc est sélectionné dans le groupe constitué par les séquences SEQ ID NO:62-67 et 119 et leurs fragments correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires.

Selon un autre mode de mise en œuvre avantageux de ladite méthode, ledit support comprend en outre toute séquence rétrovirale endogène ou  
20 exogène.

La méthode des puces à ADN (Bowtell, (1999), Nature Genet., 21, 25- 32), est utilisée pour évaluer la modification de l'expression de tout ou partie de certaines des séquences d'origine rétrovirale de type HERV-7q et des séquences flanquantes. Brièvement de l'ADN provenant de clones, de produits de PCR obtenus à  
25 partir d'ADN génomique, de produits de RT-PCR provenant de transcrits ou encore de séquences oligonucléotidiques spécifiques, sont déposées sur un support, comme par exemple un filtre de nylon, une lame de verre ou leur équivalent. Les séquences nucléiques déposées couvrent les différents domaines rétroviraux décrits ci-dessus, ainsi que les séquences contiguës et les gènes flanquants. Afin de détecter d'éventuels  
30 processus d'épissage alternatifs, des ADN spécifiques sont synthétisés par pas de 500-600 nucléotides avec un chevauchement de 250- 300 nucléotides de part et d'autre.

Les épissages alternatifs déjà identifiés feront l'objet d'une synthèse spécifique. L'hybridation s'effectue à l'aide d'une sonde obtenue, par exemple, par rétrotransposition d'un mélange d'ARN provenant de cellules, de tissus ou de liquides biologiques provenant de témoins réputés normaux, de membres de populations ethniques différentes, de patients atteints de pathologies souvent associées à une expression de rétrovirus, comme les processus tumoraux, ou comme les maladies auto-immunes, dont la sclérose en plaques. Dans ce cas une fraction de  $\mu\text{g}$  et jusqu'à quelques  $\mu\text{g}$  d'ARNm ou jusqu'à quelques  $\mu\text{g}$  ou dizaines de  $\mu\text{g}$  d'ARN, selon la méthode utilisée et la taille de la puce d'ADN concernée, sont suffisants pour la synthèse de la sonde nucléique.

10 La sonde nucléique est marquée de manière adéquate, afin d'autoriser une détection ultérieure, comme par exemple par fluorescence ou par une méthode équivalente.

L'usage de sondes bi-, voire multicolores permet de préciser l'expression concertée de plusieurs gènes en parallèle, en bénéficiant de plus d'une normalisation précise. L'acquisition des résultats est effectuée automatiquement,

15 comme par exemple par un système de balayage laser ou son équivalent.

Deux types de puces à ADN sont conçues, d'une part des puces présentant un ensemble exhaustif de séquences, et d'autre part des puces à ADN spécifiques permettant un ciblage sur une application plus spécifique.

Par exemple, une séquence critique en ce qu'elle contiendrait une

20 différence portant sur une délétion, voire une mutation, est détectée à l'aide d'oligonucléotides spécifiques (Wang et coll., (1998), Science, 280, 1077- 1082). Le polymorphisme associé à une base ou à une mutation est détecté à l'aide de quatre oligonucléotides possédant une des quatre possibilités de séquence au niveau d'une base (A, C G ou T): pour chaque différence ponctuelle les 4 oligonucléotides sont

25 déposés et les intensités d'hybridation sont comparées. De plus, un épissage alternatif est détecté en utilisant des ADN correspondant à un seul exon effectif ou putatif: le gène est donc analysé exon par exon. Les puces à ADN concernent aussi, par extension, toute séquence rétrovirale endogène ou exogène, comme par exemple ERV-9, ERV-K, ERV-L, ERV-H, ERV-4, ERV-6, ERV-8, ERV-10, ERV-15, ERV-16, ERV-

30 17, ERV-18, ERV-21, ERV-24, ERV-33, ERV-34, ERV-36, ERV-40, ERV-42, ERV-MLN, ERV-FRD, ERV-FTD...), ainsi que toutes les séquences exoniques putatives

(identifiées par l'existence de séquences étiquettes et des transcrits correspondants) ou effectives, et qui sont situées de part et d'autre jusqu'à une distance de 120 kb des séquences rétrovirales endogènes de type HERV-7q.

L'étude comparative est menée entre un échantillon témoin et  
5 l'échantillon à tester, dans une perspective prophylactique, diagnostique ou thérapeutique, comme par exemple: la détection précoce d'une modification de l'expression d'une des séquences, dans une cellule, un tissu ou un organisme, l'identification d'une séquence associée à une susceptibilité ou à une pathologie quelconque, le suivi de l'évolution de la pathologie, ou encore le suivi d'un traitement et l'évaluation de son  
10 efficacité.

En dehors des applications déjà énoncées, l'intérêt de la méthode permet, d'une manière plus générale, de faire un bilan des variations constatées chez un individu, ce qui constitue en quelque sorte une carte d'identité, ce qui facilite une approche épidémiologique permettant d'établir de nouvelles corrélations entre un  
15 profil particulier observé et une pathologie, en dehors de tout *a priori* concernant cette pathologie.

La présente invention a également pour objet un kit de détection et/ou d'évaluation d'une maladie auto-immune et notamment des neuropathologies à étiologie auto-immune, caractérisé en ce qu'il comprend outre les tampons nécessaires  
20 à la mise en œuvre des procédés tels que définis ci-dessus :

- des réactifs A de diagnostic tels que définis ci-dessus, et
- des réactifs B constitués par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou l'altération est associée à l'expression normale ou pathologique  
25 ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences correspondant à des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite famille HERV-7q dont l'une des extrémités ne peut se trouver à une distance excédant 120 kb,

30 lesquels réactifs sont de préférence fixés sur un support approprié.

Selon un mode de réalisation avantageux dudit kit, lesdits réactifs B



sont sélectionnés dans le groupe constitué par les séquences SEQ ID NO:62-67 et 119 et leurs fragments correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires, ainsi que les séquences représentées aux figures 13-17, 22-26.

5 La présente invention a également pour objet des produits de traduction, caractérisés en ce qu'ils sont codés par une séquence nucléotidique telle que définie ci-dessus.

La présente invention a également pour objet un peptide, caractérisé en ce qu'il est susceptible d'être exprimé à l'aide d'une séquence nucléotidique sélectionnée dans le groupe constitué par les séquences SEQ ID NO:1-22, 28 et 61, telles  
10 que définies ci-dessus, selon les combinaisons offertes par l'usage des différents cadres de lecture possibles (voir également figures 18-21).

Ledit peptide englobe également les peptides ou polypeptides dérivés comprenant entre 5 et 540 aminoacides (SEQ ID NO:23-36 et SEQ ID NO:58 et  
15 leurs fragments d'au moins 5 aminoacides) et notamment un fragment de 538 aminoacides, commençant à la première méthionine de la séquence SEQ ID NO:26 (envérine).

Selon un mode de réalisation avantageux desdits peptides, ils sont notamment sélectionnés parmi les séquences SEQ ID NO:23-36, 58, notamment la  
20 séquence SEQ ID NO:26 et ses fragments C-terminaux, soit à partir de l'acide aminé 291, soit à partir de l'acide aminé 321, à compter de la première méthionine.

Selon un autre mode de réalisation avantageux desdits peptides, ils sont obtenus à partir des séquences nucléiques telles que définies ci-dessus, dans lesquelles au moins un codon non-sens peut être remplacé par un codon codant pour  
25 l'un des aminoacides suivants : Phe (F), Leu (L), Ser (S), Tyr (Y), Cys (C), Trp (W), Gln (Q), Arg (R), Lys (K), Glu (E) ou Gly (G).

L'invention englobe ainsi les peptides déduits ou les protéines déduites correspondant à tout ou partie des séquences nucléiques décrites dans l'invention, et présentant éventuellement des modifications avec les séquences de réfé-  
30 rences décrites dans l'invention, lorsqu'ils sont exprimés chez certains patients. En particulier, l'invention englobe les séquences complètes ou partielles obtenues selon

les 3 cadres de lecture sens et les 3 cadres de lecture inverses et complémentaires.  
(voir figures 18-21)

De manière avantageuse, l'analyse de la structure du domaine env de  
HERV-7q, appelé envérine, a permis de mettre successivement en évidence:

- 5                   - un peptide signal N-terminal (région 1- 21) et deux domaines  
transmembranaires (région 320-340; 455-477), responsables d'interactions avec des  
motifs protéiques ou lipidiques membranaires,
  - un motif immunomodulateur de type CKS-17(Haraguchi et coll.,  
(1995), 92, 5568- 5571)/ CKS-25. On peut noter à cet égard, la présence d'un motif
- 10   **RalD** à l'intérieur du peptide de type CKS-17/CKS-25 de HERV-7q et un motif **RvaD**  
en position 363 qui correspondent au consensus W/RxxD, proposé pour le site actif  
des TGF- $\beta$  (Huang et al., J. Biol. Chem., 1997, 272, 27155- 27159). de puissants  
facteurs associés à la croissance, à la différenciation et à la morphogenèse et qui sont  
associés à de nombreuses pathologies humaines, comme les processus tumoraux
- 15 (Tang et coll., (1998), Nat. Med., 4, 802- 807) ou les maladies neurodégénératives  
(Flanders et coll., (1998), Prog. Neurobiol., 54, 71- 85). Les peptides selon l'invention  
contenant ces motifs peuvent avantageusement servir d'antagonistes en inhibant la  
fixation des TGF-  $\beta$  sur leurs récepteurs naturels.
  - des motifs de N-glycosylation. La glycosylation des protéines
- 20 d'enveloppe des rétrovirus semble être directement associée à leurs propriétés fonc-  
tionnelles, par exemple en influençant le nombre des déterminants disponibles dans les  
cellules-T ou en favorisant la reconnaissance des antigènes par les cellules-T. La  
glycosylation pourrait jouer un rôle dans la déclaration ou l'extension d'une patholo-  
gie à incidence autoimmune. Les glycosylations sont nécessaires au maintien de la
- 25 conformation de certains épitopes, en particulier lors de la réalisation d'une protéine  
d'enveloppe recombinante à fin de mise au point d'un réactif de diagnostic et pour  
favoriser l'efficacité d'un éventuel vaccin. Positions 171, 210, 216, 236, 244, 283 et  
411. Nombre prévu au hasard : 3.2
  - des sites de prénylation. La prénylation est un mécanisme essentiel
- 30 de la fixation à la membrane cellulaire et pour le ciblage de certaines protéines. Ce  
processus de ciblage pourrait être essentiel pour l'élaboration d'agents thérapeutiques

spécifiques aptes à interférer dans la réalisation et la régulation du trafic de complexes cellulaires mettant en jeu des protéines impliquées dans les interactions, la croissance et les mouvements cellulaires. Positions 188 et 290. Nombre prévu au hasard : 1.8

- des sites de ciblage dans le réticulum endoplasmique. Ces sites
- 5 permettraient d'assurer le ciblage vers le réticulum endoplasmique afin d'effectuer les modifications nécessaires pour favoriser le franchissement membranaire. Positions 353 et 431. Nombre prévu au hasard : 0.2.

Par ailleurs, les Inventeurs ont montré qu'un certain nombre de peptides issus de la protéine env de HERV-7q (envérine) présentent une affinité/demi-  
 10 vie élevées pour des allèles HLA de classe I. Une analyse par CADD a permis de sélectionner des peptides candidats, dont les meilleurs scores sont indiqués dans le Tableau I:

TABLEAU I

15	locali- sation	séquence	molécule HLA	score	Séquence n°
		399 FLGEECCYYV	A-0201	7214	SEQ ID NO:68
		462 LLFGPCIFNL	A-0201	1792	SEQ ID NO:69
20		189 CLPLNFRPYV	A-0201	1453	SEQ ID NO:70
		439 GLLSQWMPWI	A-0201	488	SEQ ID NO:71
		263 CLPSGIFV	A-0201	5103	SEQ ID NO:72
		444 WMPWILPFL	A-0201	897	SEQ ID NO:73
		252 IRWVTPPTQI	B-2705	3000	SEQ ID NO:74
25		432 LRNTGPWGLL	B-2705	2000	SEQ ID NO:75
		158 LRTHTRLVSL	B-2705	2000	SEQ ID NO:76
		316 KRVPILPFVI	B-2705	1800	SEQ ID NO:77
		25 CRCMTSSSPY	B-2705	1000	SEQ ID NO:78
		137 TRVHGTSSPY	B-2705	1000	SEQ ID NO:79
30		124 AREKHVKEVI	B-2705	600	SEQ ID NO:80
		478 SRIEAVKLQM	B-2705	600	SEQ ID NO:81
		442 SQWMPWILPF	B-2705	500	SEQ ID NO:82
		405 CYYVNQSGI	Kd	2400	SEQ ID NO:83
		346 FYYKLSQEL	Kd	2400	SEQ ID NO:84
35		244 TYTTNSQCI	Kd	2400	SEQ ID NO:85
		291 SFLVPPMTI	Kd	1600	SEQ ID NO:86
		406 YYVNQSGIV	Kd	1200	SEQ ID NO:87
		167 LFNTTLTGL	Kd	1152	SEQ ID NO:88
		463 LFGPCIFNL	Kd	960	SEQ ID NO:89
40		253 RWVTPPTQI	Kd	480	SEQ ID NO:90
		449 LPFLGPLAAI	B-5102	2200	SEQ ID NO:91
		3 LPYHIFLFTV	B-5102	1210	SEQ ID NO:92

TABLEAU I (suite)

	locali- sation	séquence	molécule HLA	score	Séquence n°
5					
	331	GALGTGIGGI	B-5102	798	SEQ ID NO:93
	321	LPFVIGAGVL	B-5102	550	SEQ ID NO:94
	499	RRPLDRPAS	B-2705	600	SEQ ID NO:95
10	194	FRPYVSIPV	B-2705	600	SEQ ID NO:96
	383	RRALDLLTA	B-2705	600	SEQ ID NO:97
	39	WRMQRPGNI	B-2705	600	SEQ ID NO:98
	423	DRIQRAEEL	B14	1800	SEQ ID NO:99
	158	LRTHTRLVSL	B14	600	SEQ ID NO:100
15	359	ERVADSLVTL	B14	540	SEQ ID NO:101
	463	LFGPCIFNLL	Kd	1658	SEQ ID NO:102
	345	QFYKLSQEL	Kd	1152	SEQ ID NO:103
	443	QWMPWILPFL	Kd	691	SEQ ID NO:104
	405	CYYVNQSGIV	Kd	500	SEQ ID NO:105
20	474	NFVSSRIEAV	Kd	480	SEQ ID NO:106
	221	GPLVSNLEI	B-5102	1320	SEQ ID NO:107
	190	LPLNFRPYV	B-5102	726	SEQ ID NO:108
	449	LPFLGPLAAI	B-5101	1144	SEQ ID NO:109
	488	EPKMQSCKTI	B-5101	968	SEQ ID NO:110
25	3	LPYHIFLFTV	B-5101	629	SEQ ID NO:111
	125	REKHVKEVI	Kk	1000	SEQ ID NO:112
	312	KPRNKRVPIL	B7	800	SEQ ID NO:113
	378	VVLQNRRAI	Db	792	SEQ ID NO:114
	377	AVVLQNRRAI	Db	660	SEQ ID NO:115
30	321	LPFVIGAGV	B-5101	629	SEQ ID NO:116
	304	DLYSYVISK	A3	540	SEQ ID NO:117
	301	TEQDLYSYVI	Kk	500	SEQ ID NO:118

Ce Tableau I indique une estimation de la demi-vie de dissociation d'un peptide de l'envérine avec un allèle du système HLA de classe I (les tables de coefficients de Parker: J. Immunol, (1994), 152, 163- 175). La localisation indique la position du premier acide aminé des peptides testés dans la séquence de l'envérine. Le code à une lettre est utilisé pour la séquence des acides aminés. Les scores autour de 500 ou supérieurs à 500 ont été retenus. A titre de comparaison, une analyse a été effectuée sur une concaténation de peptides (polypeptide de 4968 acides aminés) réputés pour fixer les molécules du complexe majeur d'histocompatibilité de classe I (Rammensee, Immunogenetics, (1995), 41, 178- 228): les dix meilleurs scores enregistrés pour des nonapeptides et le type HLA, A\_0201 sont respectivement de 4984,

4047, 2406, 1267, 800, 705, 607, 591, 591 et 577.

Il ressort de ce Tableau I que certaines molécules du complexe majeur d'histocompatibilité de type I sont aptes à fixer des peptides issus de l'environnement, ainsi assimilés à des peptides d'origine virale ou tumorale, au niveau du réticulum endoplasmique. Les complexes formés au niveau du réticulum endoplasmique sont alors transportés à la surface cellulaire, ce qui entraîne la destruction de la cellule cible par les lymphocytes-T cytotoxiques. Les peptides identifiés comportent généralement 8 à 10 acides aminés. Des études ont montré que certains allèles du système HLA de classe I sont ainsi associées à certaines pathologies, en particulier à caractère auto-immun, comme HLA-B27 avec la spondylarthrite ankylosante ou HLA-B51 avec la maladie de Behçet.

Un peptide apte à fixer une molécule particulière de classe I est par conséquent apte à fonctionner comme un épitope de cellule-T.

En conséquence, la présente invention inclut également les fragments 399-471 et 244-271 de l'environnement qui regroupent avantageusement plusieurs épitopes de forte affinité pour différents haplotypes du système HLA de classe I. L'usage de tout ou partie de ces polypeptides est en conséquence apte à favoriser une augmentation du répertoire des cellules-T, en permettant une meilleure efficacité de la réponse immunitaire dans le cadre des différentes stratégies immunothérapeutiques, prophylactique ou vaccinales). Ces peptides pourront être avantageusement délivrés par exemple par l'usage, de vecteurs viraux, de particules virales ou synthétiques, de lipopeptides, d'adjuvants classiques, d'acides nucléiques nus ou adsorbés sur des particules, ou de liposomes.

Au sens de la présente invention, les peptides peuvent être chimiquement ou biochimiquement modifiés ; certaines des acides aminés peuvent être remplacés par un acide aminé analogue, selon les critères classiques d'homologies (A ou G ; S ou T ; I, L ou V ; F, Y ou W ; N ou Q ; D ou E).

La présente invention a également pour objet des compositions immunogènes ou vaccinales, pour la protection contre les maladies auto-immunes, notamment chez les sujets à risque, caractérisée en ce qu'elle comprend au moins un peptide comprenant au moins un motif de type CKS et/ou au moins un peptide

constitué par un motif présentant une affinité avec l'un des haplotypes du système HLA de classe I ou de classe II et un véhicule pharmaceutiquement acceptable.

Selon un mode de réalisation avantageux de ladite composition, ledit motif est sélectionné dans le groupe constitué par les peptides tels que définis dans le  
5 Tableau I ci-dessus.

Selon un autre mode de réalisation avantageux de ladite composition, ledit peptide présente la séquence suivante :

séquence CKH: LONRRALDLLTAERGGTcIFLGEECCYYV  
(SEQ ID NO:120).

10 Il est remarquable de constater au niveau de la position 380 de la protéine envérine, la contiguïté des motifs de type CKS-17 (souligné) et du peptide présentant le score le plus élevé (en gras ; voir peptide en position 399 dans le Tableau I, SEQ ID NO:68) dans la séquence CKH.

L'activation clonale des sous-groupes de lymphocytes, par exemple  
15 de lymphocytes cytotoxiques, par les peptides du Tableau I et par extension leurs homologues, est bloquée par des manœuvres usuelles d'immunothérapie comme par exemple la sérothérapie et la vaccination.

L'association de deux séquences ou des séquences analogues au peptide CKH (SEQ ID NO:120), est à même d'entraîner un processus synergique dans  
20 la réponse immunitaire, qui pourrait mettre en jeu des voies de signalisation et d'activation complémentaires, aptes à moduler l'activation lymphocytaire.

La vaccination concerne la production d'anticorps dirigés contre les peptides du tableau I, selon les règles de l'art et selon les méthodes de libération contrôlées par implants artificiels ou cellulaires mettant en œuvre une composition  
25 telle que définie ci-dessus et par usage des manœuvres de thérapie génique, comme par exemple par expression des séquences nucléiques codant pour les peptides du Tableau I. En conséquence l'invention a également pour objet des compositions immunogènes ou vaccinale caractérisée en ce qu'elles comprennent un vecteur incluant au moins une séquence nucléique codant un peptide tel que défini dans le  
30 Tableau I, éventuellement associée à une séquence codant un motif de type CKS-17.

La sérothérapie concerne l'utilisation d'anticorps neutralisants

produits à partir des peptides du Tableau I et leurs homologues.

Les produits protéiques générés par les séquences rétrovirales endogènes ou produits parallèlement peuvent avantageusement être caractérisés par des micro-méthodes d'analyse et de quantification des peptides et des protéines:

5 HPLC/FPLC ou équivalent, électrophorèse capillaire ou équivalent, techniques de microséquençages (méthode d'Edman ou équivalent, spectrométrie de masse...).

L'invention a également pour objet des anticorps dirigés contre l'un ou plusieurs des peptides décrits ci-dessus et leur utilisation soit pour la mise en œuvre d'une méthode de détection *in vitro*, notamment différentielle de la présence d'une

10 telle séquence chez un individu, soit pour la préparation d'une composition apte à être utilisée en sérothérapie dans les neuropathologies à composante auto-immune.

Lesdits anticorps sont avantageusement des anticorps polyclonaux ou monoclonaux obtenus par une réaction immunologique d'un organisme humain, mammifères, oiseaux ou autres espèces vis-à-vis des protéines, telles que définies ci-

15 dessus.

La présente invention a pour objet un procédé de dépistage immunologique différentiel de séquences rétrovirales endogènes humaines de la famille HERV-7q normales ou pathologiques, caractérisé en ce qu'il comprend la mise en contact d'un échantillon biologique avec un anticorps selon l'invention, la lecture

20 du résultat étant révélée par un moyen approprié, notamment EIA, ELISA, RIA, fluorescence.

A titre d'illustration, une telle méthode de diagnostic *in vitro* selon l'invention comprend la mise en contact d'un échantillon biologique prélevé chez un patient, avec des anticorps selon l'invention et la détection à l'aide de tout procédé

25 approprié, notamment à l'aide d'anti-immunoglobulines marquée, des complexes immunologiques formés entre les protéines produites normalement ou pathologiquement et les anticorps.

Des anticorps monoclonaux ou polyclonaux, produits à partir d'antigènes correspondants à des peptides de synthèse, de polypeptide ou protéines

30 recombinants, permettent de suivre l'expression des peptides ou protéines produits normalement ou pathologiquement. L'analyse est de préférence effectuée par ELISA,

ou équivalent, Western-blot ou équivalent, ou par immunohistochimie.

Les peptides ou protéines, issus des séquences rétrovirales endogènes ou dont l'expression est associée à l'expression de ces séquences rétrovirales endogènes, sont recherchés et identifiés.

5 La présente invention a également pour objet un procédé d'identification et de détection de motifs rétroviraux endogènes, anormalement exprimés dans le cadre de pathologies associées au cancer, ou de neuropathologies en particulier auto-immunes, au premier rang desquelles la sclérose en plaques, caractérisé en ce qu'il comprend l'analyse comparée des séquences extraites d'un échantillon biologique  
10 avec les séquences selon l'invention.

La présente invention a également pour objet l'application des séquences nucléiques ou des séquences protéiques selon l'invention au diagnostic, au pronostic, à l'évaluation de la susceptibilité génétique, à toutes maladies humaines induites, innées ou acquises en particulier celles à composantes cancéreuses, auto-  
15 immunes et/ou à incidence neurologique, comme la sclérose en plaques, les syndromes associés et les maladies neurodégénératives où intervient tout ou partie des séquences nucléiques selon l'invention et des formes endogènes ou exogènes apparentées.

La présente invention a également pour objet des séquences  
20 nucléiques hybrides, caractérisées en ce qu'elles comprennent des séquences ou motifs nucléiques selon l'invention, combinés avec des séquences ou motifs d'origine endogène ou d'origine ou induits de manière exogène.

La présente invention a, en outre, pour objet un vecteur recombinant de clonage ou d'expression, caractérisé en ce qu'il comprend une séquence nucléique  
25 conforme à l'invention.

Des manœuvres thérapeutiques peuvent être envisagées par usage de certaines des séquences nucléiques contenues dans HERV-7q et les séquences de la même famille ou des structures polypeptidiques déduites ou par utilisation de peptides ou protéines, ou d'anticorps spécifiques.

30 Conformément à l'invention, tout ou partie des séquences nucléiques rétrovirales endogènes de type HERV-7q, peut être utilisée à usage de vecteur ou



d'éléments de vecteurs à vocation thérapeutique, en particulier les séquences LTR et la région *gag* (SEQ ID NO :2, 21 et 22)..

L'intérêt de telles séquences réside, dans l'innocuité du vecteur ainsi formé, dans la possibilité d'une insertion spécifique ciblée dans une région bien définie par une stratégie analogue à la recombinaison homologue, dans le ciblage cellulaire, éventuellement transitoire dans le cas d'une expression placentaire chez la femme. Un autre aspect concerne la possibilité d'associer aux gènes d'intérêts les motifs rétroviraux biologiquement actifs (peptides immunomodulateurs, tels que représentés aux séquences SEQ ID NO 68-118, ci-après, peptide fusogène...).

La présente invention a également pour objet des animaux transgéniques, caractérisés en ce qu'ils comprennent tout ou partie d'une séquence de type HERV-7q (SEQ ID NO:1-22 et 61).

Le Tableau II ci-après établit les correspondances entre les numéros des séquences telles qu'elles apparaissent dans la liste de séquences et le nom des différentes séquences.

TABLEAU II

SEQ ID NO :	DESIGNATION
1	Acide nucléique : 7 env
2	Acide nucléique : gag
3	Acide nucléique : HERV-7q
4	Acide nucléique : HE2
5	Acide nucléique : HE3
6	Acide nucléique : HG3
7	Acide nucléique : HE4
8	Acide nucléique : HE5
9	Acide nucléique : HE6
10	Acide nucléique : HG6
11	Acide nucléique : HE7
12	Acide nucléique : HE8
13	Acide nucléique : HG8
14	Acide nucléique : HE9
15	Acide nucléique : HE10
16	Acide nucléique : HE11
17	Acide nucléique : HG11
18	Acide nucléique : HE12
19	Acide nucléique : HG12

SEQ ID NO:	DÉSIGNATION
20	Acide nucléique : R1
21	Acide nucléique : R1F
22	Acide nucléique + protéine env déduite : HERV-7q
23	Fragment de protéine env déduite selon SEQ ID NO :22
24	Fragment de protéine env déduite selon SEQ ID NO :22
25	Fragment de protéine env déduite selon SEQ ID NO :22
26	Protéine : envérine
27	Fragment de protéine env déduite selon SEQ ID NO :22
28	Acide nucléique + protéine déduite de gag : HERV-7q
29	Fragment de protéine gag déduite selon SEQ ID NO :28
30	Fragment de protéine gag déduite selon SEQ ID NO :28
31	Fragment de protéine gag déduite selon SEQ ID NO :28
32	Fragment de protéine gag déduite selon SEQ ID NO :28
33	Fragment de protéine gag déduite selon SEQ ID NO :28
34	Fragment de protéine gag déduite selon SEQ ID NO :28
35	Protéine env : cadre de lecture 1
36	Protéine gag
37	Acide nucléique : G1F (amorce)
38	Acide nucléique : G1R (amorce)
39	Acide nucléique : G2F (amorce)
40	Acide nucléique : G2R (amorce)
41	Acide nucléique : G4F (amorce)
42	Acide nucléique : G3F (amorce)
43	Acide nucléique : G4R (amorce)
44	Acide nucléique : G5R (amorce)
45	Acide nucléique : E1F (amorce)
46	Acide nucléique : E1R (amorce)
47	Acide nucléique : E2F (amorce)
48	Acide nucléique : E2R (amorce)
49	Acide nucléique : E3F (amorce)
50	Acide nucléique : E3R (amorce)
51	Acide nucléique : E4F (amorce)
52	Acide nucléique : E4R (amorce)
53	Acide nucléique : E5F (amorce)
54	Acide nucléique : E6F (amorce)
55	Acide nucléique : E5R (amorce)
56	Acide nucléique : ExF (amorce)
57	Acide nucléique : ExR (amorce)
58	Protéine gag
59	Acide nucléique : Séquence A (séquence d'insertion)
60	Acide nucléique : Séquence B (séquence d'insertion)
61	Acide nucléique : HE13
62	Acide nucléique : RH7

SEQ ID NO:	DÉSIGNATION
63	Acide nucléique : RAM75
64	Acide nucléique : RAV73
65	Acide nucléique : RBP3
66	Acide nucléique : HI3
67	Acide nucléique : LTX
68	Peptide Tableau I
69	Peptide Tableau I
70	Peptide Tableau I
71	Peptide Tableau I
72	Peptide Tableau I
73	Peptide Tableau I
74	Peptide Tableau I
75	Peptide Tableau I
76	Peptide Tableau I
77	Peptide Tableau I
78	Peptide Tableau I
79	Peptide Tableau I
80	Peptide Tableau I
81	Peptide Tableau I
82	Peptide Tableau I
83	Peptide Tableau I
84	Peptide Tableau I
85	Peptide Tableau I
86	Peptide Tableau I
87	Peptide Tableau I
88	Peptide Tableau I
89	Peptide Tableau I
90	Peptide Tableau I
91	Peptide Tableau I
92	Peptide Tableau I
93	Peptide Tableau I
94	Peptide Tableau I
95	Peptide Tableau I
96	Peptide Tableau I
97	Peptide Tableau I
98	Peptide Tableau I
99	Peptide Tableau I
100	Peptide Tableau I
101	Peptide Tableau I
102	Peptide Tableau I
103	Peptide Tableau I
104	Peptide Tableau I
105	Peptide Tableau I

106	Peptide Tableau I
SEQ ID NO:	DÉSIGNATION
107	Peptide Tableau I
108	Peptide Tableau I
109	Peptide Tableau I
110	Peptide Tableau I
111	Peptide Tableau I
112	Peptide Tableau I
113	Peptide Tableau I
114	Peptide Tableau I
115	Peptide Tableau I
116	Peptide Tableau I
117	Peptide Tableau I
118	Peptide Tableau I
119	Acide nucléique : BLIMP-1
120	Peptide : CKH
121	Acide nucléique : F645 (amorce)
122	Acide nucléique : PS5D (amorce)

Outre les dispositions qui précèdent, l'invention comprend encore d'autres dispositions, qui ressortiront de la description qui va suivre, qui se réfère à des exemples de mise en œuvre du procédé objet de la présente invention ainsi qu'aux

5 dessins annexés, dans lesquels :

- Figure 1. Séquence nucléique humaine HERV-7q, dont l'analyse et le traitement permettent de caractériser une nouvelle structure rétrovirale endogène. Les régions nucléiques répétées de type R1 et R2 et les domaines *gag*, *pol* et *env* sont soulignés. Les domaines de type *gag* et *env* sont en italiques. La région homologe à une partie 3' non-codante de Rab7 est doublement soulignée.

- Figure 2. Cartographie de la région rétrovirale endogène humaine HERV-7q. La partie haute de la figure correspond à une région anonyme du génome humain située sur le bras long du chromosome 7. On peut identifier les domaines répétés (1), *gag* (2), *pol* (3) et *env* (4) de HERV-7q. La région *env* C-terminale (4.3) se prolonge en amont en un long cadre de lecture ouvert (4.2). Le domaine 4.1, correspond à la région N-terminale du domaine *env*.

- Figure 3. Comparaison des séquences nucléiques répétées situées aux bornes de HERV-7q. Les régions nucléiques répétées 5'(haut) et 3'(bas), sont

comparées et les bases identiques sont indiquées par deux points.

- Figure 4. Séquence déduite présentant un cadre de lecture ouvert, dans le domaine de type-env de HERV-7q selon la règle du plus long cadre de lecture ouvert.

5                   - Figure 5. Séquences autour du domaine CKS-17 identifiées dans différents domaines *env* déduits de la famille de HERV-7q et comparaison avec des motifs CKS-17 de référence.

1) HE2 - 2) HERV-7q - 3) N° d'accès à GenBank: M85205 - 4) HE7 - 5) HE9 - 6) CKS-17: le motif peptidique doué de propriétés immunomodula-  
10 trices est souligné - 7) gp20 de rétrovirus de type-D (SRV-Pc).

- Figure 6. Séquence déduite possible du domaine de type-*gag* identifié dans HERV-7q établie selon la règle du plus long cadre de lecture ouvert. X et / correspondent respectivement à un codon non-sens et à un décalage de cadre de lecture. La séquence soulignée correspond au début du domaine *pol*.

15                   - Figure 7. Comparaison des régions nucléiques couvrant la région *gag* de HERV-7q (haut) et HERV-TcR (bas) et leurs régions flanquantes. Les bases identiques sont spécifiées par deux points.

- Figure 8. Exemple d'alignements nucléiques du domaine de type *env* de HERV-7q avec des domaines de type *env* similaires présents dans des  
20 séquences rétrovirales endogènes humaines de la même famille. Les codons non sens sont soulignés : 1) HERV-7q - 2) HE2 - 03) HE3 - 04) HE4.

- Figure 9. Alignements nucléiques entre le domaine *gag* de HERV-7q et les domaines correspondants appartenant à la même famille. Comparaison avec des fragments de domaines *gag* isolés d'agents rétroviraux infectieux. Séquences  
25 d'origine rétrovirale infectieuse: N° d'accession dans la banque de données EMBL : 1) A60168 - 2) A60201 - 3) A60200 - 4) A60171. Séquences rétrovirales endogènes humaines: 5) HERV-7q - 6) HG11 - 7) HG3. Les chiffres indiqués dans les séquences endogènes, correspondent au nombre de nucléotides insérés afin d'optimiser l'alignement avec les séquences de type *gag* identifiées dans des rétrovirus d'origine  
30 infectieuse.

- Figure 10. Alignement d'un motif *gag* protéique déduit (haut)

appartenant à un rétrovirus infectieux (N° d'accèsion EMBL : A60200) avec le motif *gag* protéique déduit (bas) identifié dans HERV-7q. Les codons non-sens sont en gras et soulignés. Les acides aminés identiques sont spécifiés par 2 tirets. Un tiret indique une délétion ou un acide aminé homologue.

5                   - Figure 11. Alignement d'un motif *env* (haut) appartenant à un rétrovirus infectieux (N° d'accèsion EMBL : A60170) avec le motif *env* (bas) identifié dans HERV-7q. Les nucléotides homologues sont spécifiés par deux points et les délétions par un tiret.

10                   - Figure 12. Comparaison entre le domaine *env* de HERV-7q (haut) et le domaine *env* de HERV-9 (bas). L'homologie de 66 % se limite à la région 3' du domaine *env* de HERV-7q et HERV-9, respectivement entre les nucléotides 8976 nt et 9500 nt de HERV-7q et les nucléotides 2898 nt et 3465 nt de HERV-9 (N° d'accèsion à GenBank : X57147). De nombreuses insertions/délétions sont aussi observées.

15                   - Figure 13. Homologie entre une partie de la séquence du transcrit codant pour RH7 (haut, SEQ ID NO:62) et un motif de RGH2 (bas - N° d'accèsion à GenBank: D11018).

20                   - Figure 14. Identification de la séquence du transcrit codant pour RAM75 (SEQ ID NO:63), correspondant au gène d'une ATPase de type PEX1. Les exons codants sont soulignés. Les codons d'initiation et non-sens ainsi que les sites putatifs de polyadénylation sont en gras et soulignés. La région en italique correspond au début de la séquence rétrovirale endogène RH7.

25                   - Figure 15. Séquence du transcrit codant pour RAV73 (SEQ ID NO:64), située à 0.7 kb en aval de HERV-7q ; les séquences nucléiques aptes à coder pour un ou plusieurs polypeptides sont soulignées.

                  - Figure 16. Comparaison entre la séquence LTR 3' (haut) de HERV-7q et la séquence intronique LTX (SEQ ID NO:67), située dans le gène FMR2, associé au X-fragile (bas).

30                   - Figure 17. Mise en évidence de modifications sur la séquence nucléotidique (ID NO:3), chez des patients atteints de SEP. Les bases modifiées, chez au moins un patient, sont soulignées. Les amorces utilisées sont en italiques

(séquences SEQ ID NO:121 et 122). L'ATG d'initiation et le codon non-sens sont en gras.

- Figure 18. Partie codante *env* de la séquence HERV-7q (séquence ID NO:3), avec 3 cadres de lecture.

5                   - Figures 19, 20, 21. Présentation séparée de la protéine *env* selon les 3 cadres de lecture.

                  - Figure 22. Séquence nucléique contenant la séquence rétrovirale RH7 située en 5' de la séquence HERV-7q. La séquence en italique correspond au début de la séquence HERV-7q. La séquence RH7 est soulignée. Deux sites de poly-  
10 adénylation putatifs sont gras.

                  - Figure 23. Séquence du transcrit codant pour RBP3 contenant des motifs nucléotidiques identifiés dans la séquence nucléique codant pour le gène Blimp-1.

                  - Figure 24. Séquence du transcrit codant pour APS.

15                   - Figure 25. Séquence du transcrit codant pour Blimp-1 ; la partie codante est soulignée ; les codons d'initiation et de terminaison sont en gras.

                  - Figure 26. Séquence du transcrit codant pour FMR2. La partie codante est soulignée. Les codons d'initiation et non-sens sont en gras.

                  Il doit être bien entendu, toutefois, que ces exemples sont donnés  
20 uniquement à titre d'illustration de l'objet de l'invention, dont ils ne constituent en aucune manière une limitation.

**EXEMPLE 1 : Détection, par amplification génique, d'une séquence nucléique appartenant à un domaine de type *gag* ou *env* selon l'invention, dans un échantillon d'ADN génomique d'origine humaine ou de mammifères.**

25                   L'amplification génique s'effectue à partir d'ADN génomique isolé à partir du sang. Un traitement anticoagulant est effectué avec 1 ml d'une solution de citrate (pour un litre : 4,8 g de d'acide citrique, 13,2 g de citrate de sodium, 14,7 g de glucose) pour 6 ml de sang frais. Après centrifugation de 20 ml de sang pendant 15 mn à 13.0000 g, le surnageant est éliminé et la fraction enrichie en globules blancs est

transférée dans un nouveau tube, puis recentrifugée dans les mêmes conditions que précédemment. La fraction enrichie en globules blancs est resuspendue dans un tampon d'extraction (10 mM Tris-HCl, 0,1 M EDTA, 20 µg/ml de RNase pancréatique traitée afin d'éliminer les DNases, 0,5 % SDS, pH 8,0), puis incubée pendant 1 heure à 37°C. La protéinase K est ajoutée à une concentration finale de 100 µg/ml. La suspension des cellules lysées est incubée à 50°C durant 3 heures sous agitation périodique, puis traitée par un volume égal de phénol équilibré par du Tris-HCl 0,5 M, pH 8,0. L'émulsion formée est placée sur une roue pendant une heure, puis centrifugée à 5000 g pendant 15 mn à température ambiante. La solution aqueuse est traitée déprotéinisée par une triple extraction phénolique afin d'obtenir un niveau de purification correspondant à un rapport final d'absorbance A260/A280 supérieur à 1,75. La fraction aqueuse est précipitée par 0,2 vol. d'acétate de sodium 10 M et 2 vol. d'éthanol. L'ADN est alors soit prélevé avec l'extrémité d'une pipette pasteur recourbée, soit centrifugé à 5000 g pendant 5 mn à température ambiante. L'ADN ou le culot d'ADN est lavé deux fois par de l'éthanol à 70 %, puis repris dans 1 ml de TE pH 8,0 afin d'être élué sous agitation douce pendant 12 à 24 heures.

Des oligonucléotides spécifiques des séquences endogènes décrites selon l'invention sont choisis pour amplifier la région *gag* ou *env* des régions rétrovirales endogènes décrites selon l'invention. L'ADN génomique étudié provient de patients présentant des pathologies comme la sclérose en plaques et d'individus réputés sains.

Les ADN polymérases thermostables utilisées ont été choisies pour leur grande fidélité lors du processus d'amplification, comme la Vent, ADN polymérase (Biolabs) ou équivalent, et sont utilisées selon les conditions préconisées par le fournisseur.

La stratégie d'amplification utilise selon les cas une simple PCR, ou une PCR nichée ou semi-nichée.

Oligonucléotides utilisés pour amplifier la région *gag* :

- amorce G1F, sens, localisée dans la région amont du domaine *gag* de *HERV-7q* (SEQ ID NO:37),
- amorce G1R, anti-sens, localisée dans la région 3' terminale du



domaine *gag* (SEQ ID NO:38),

Le fragment de 1505 nt amplifié par le couple G1F-G1R : 1505 nt est utilisé afin de générer les sondes aptes à hybrider les différents produits d'amplification des PCR.

- 5
- amorce G2F, sens nichée (SEQ ID NO:39),
  - amorce G2R, anti-sens nichée (SEQ ID NO:40),
  - amorce G4F, sens nichée (SEQ ID NO:41),
  - amorce G3F, sens nichée (SEQ ID NO:42).
  - amorce G4R, anti-sens nichée (SEQ ID NO:43),
  - 10 - amorce G5R, anti-sens nichée (SEQ ID NO:44),

Oligonucléotides utilisés pour amplifier la région *env* de HERV-7q :

- amorce E1F, sens (SEQ ID NO:45),
- amorce E1R, anti-sens (SEQ ID NO:46),

- 15 Le fragment de 2529 nt amplifié par le couple d'amorces E1F-E1R, est utilisé afin de générer les sondes aptes à hybrider les différents produits d'amplification des PCR.

- amorce E2F, sens (SEQ ID NO:47),
- amorce E2R, antisens (SEQ ID NO:48),
- amorce E3F, sens (SEQ ID NO:49),
- 20 - amorce E3R, anti-sens (SEQ ID NO:50),
- amorce E4F, sens (SEQ ID NO:51),
- amorce E4R, anti-sens (SEQ ID NO:52),
- amorce E5F, sens (SEQ ID NO:53),
- amorce E6F, sens(SEQ ID NO:54)
- 25 - amorce E5R(SEQ ID NO:55).
- amorce ExF (SEQ ID NO:56)
- amorce ExR (SEQ ID NO:57)

- La PCR est réalisée à partir de 50 à 200 ng d'ADN génomique. Les conditions de PCR sont celles préconisées par le fournisseur. Les conditions cycliques
- 30 d'amplification sont réalisées dans 50 µl : une dénaturation de 94°C pendant 1 min., une hybridation de 70°C pendant 1 min., et une élongation à 72 °C pendant 1 à 2 min.,

selon les fragments amplifiés. Après 35 cycles, une réaction terminale est menée à 72°C pendant 10 min. Le séquençage automatique des échantillons amplifiés est réalisé à l'aide d'un séquenceur Applied Biosystems de type ABI 377 ou autre modèle comparable, selon les protocoles fournis par le constructeur.

- 5                    Dans le cas d'une PCR nichée ou semi-nichée, les mêmes conditions expérimentales sont utilisées, à la seule différence que l'échantillon d'ADN génomique est remplacé par 5 à 10 µl du produit d'amplification issu de la première PCR.

Deux amplifications indépendantes sont réalisées à partir du même échantillon. Une réaction de contrôle est réalisée en remplaçant l'échantillon d'ADN  
10 par de l'eau afin de détecter d'éventuels contaminants.

**EXEMPLE 2 : Détection par amplification génique d'une séquence nucléique selon l'invention dans un échantillon biologique d'ADN génomique prélevé chez des patients présentant une pathologie candidate déclarée ou la suspicion de cette pathologie.**

- 15                    Le protocole d'amplification est le même que dans l'exemple 1, mis à part l'origine de l'échantillon qui provient de patients présentant une pathologie candidate. Un échantillon d'ADN génomique réputé normal est systématiquement intégré dans l'ensemble des échantillons pathologiques amplifiés puis analysés.

Les produits de PCR sont séparés sur un gel d'agarose à 1,5 %, puis  
20 transférés en présence de soude 0,4 N sur une membrane de nylon chargé. Une hybridation est réalisée avec une sonde spécifique correspondant aux fragments de PCR amplifiés soit par les couples G1F-G1R soit par le couple E1F-E1R. La sonde est marquée par incorporation de dUTP-digoxygénine selon le protocole du fournisseur (Boehringer Mannheim). L'hybridation est effectuée dans un tampon d'hybridation  
25 (5XSSC, 50 % formamide, 0,1 % lauroyl-sarcosine, 0,02 % SDS, 2 % de réactif de blocage Boehringer) pendant une nuit à 42°C. Le Southern est lavé 2 fois 5 min. à température ambiante dans une solution de 2XSSC, 0,1% SDS. Puis un lavage à haute stringence est effectué à deux reprises pendant 15 min. à 55°C dans une solution 0,1XSSC, 0,1 % SDS. L'hybridation est révélée selon le protocole du fournisseur  
30 (Boehringer Mannheim), en présence d'un substrat chimioluminescent de la phosphatase alcaline, de type CSPD ou CDP-STAR. Le filtre est révélé après une exposition

de 15min. à 60 min.

Une analyse par SSCP ("*single strand conformation polymorphism*") permet de détecter des modifications discrètes de la séquence des fragments amplifiés par PCR. La PCR est menée en présence de dCTP marqués au P<sup>32</sup>. L'échantillon à analyser est dénaturé à 95°C pendant 10 min., en présence de tampon de charge, puis immédiatement chargé sur un gel de polyacrylamide à 10%, contenant 7.5% de glycérol. La migration s'effectue à 4°C à 8-10 W. Le gel est séché puis autoradiographié.

Les fragments de PCR susceptibles de présenter une altération de leur séquence nucléotidique sont séquencés selon l'exemple 1.

Une hybridation à l'aide d'un oligonucléotide spécifique (17 mers à 20 mers) correspondant à la région nucléotidique modifiée permet d'identifier les échantillons présentant une modification identique (méthode ASO). Brièvement le southern est hybridé avec un oligonucléotide marqué distalement soit au P<sup>32</sup>, soit en présence de digoxygénine (selon le protocole de Boehringer Mannheim) puis lavé dans des conditions stringentes à 65°C dans une solution 6XSSC, 0.05% pyrophosphate de sodium.

Par exemple, un séquençage nucléotidique automatique a été réalisé sur six fragments de PCR, provenant de 5 patients atteints de SEP et un témoin réputé normal, et qui ont été amplifiés à partir des amorces F645: CTTCAAACAACAACCAGGAGG (SEQ ID NO:121) (située à 26 nucléotides en amont de la méthionine d'initiation de l'envérine) et PSSD: TTGGGGAGGTTGGCCGACGA (SEQ ID NO:122) (située à 6 nucléotides en aval du codon non-sens de l'envérine). Des modifications de la séquence de l'envérine ont été observés sur l'ADN de certains des patients (figure 17).

**EXEMPLE 3 : Détection d'une protéine selon l'invention dans un échantillon biologique.**

- Préparation d'une fraction protéique purifiée de liquide céphalo-rachidien de patients atteints de SEP

Après un traitement à 56°C pendant 30 min, et élimination des immunoglobulines sur une colonne de protéine G HiTrap (Pharmacia), l'équivalent de

10 ml de LCR est déposé sur une colonne de DEAE Sepharose CL-6B (Pharmacia). L'élution est réalisée en Tris-HCl 20 mM pH 8,8, et un gradient de 0 à 0,4 M de NaCl, puis la fraction est dialysée 2 fois contre du tampon phosphate-NaCl (PBS). Après concentration sur Ultrafree-MC (Millipore), la fraction est déposée sur une colonne de  
 5 Superose 12 (FPLC Pharmacia) et éluée en présence de PBS. Après séparation par électrophorèse en gel de polyacrylamide-SDS, et électro-transfert sur une membrane d'Immobilon-P (Millipore), les bandes protéiques sont soumises à une hydrolyse trypsique ménagée.

- Analyse de la fraction protéique par spectrométrie de masse

10 Les peptides digérés en présence de trypsine, sont analysés par la méthode de MALDI-TOF, qui permet l'analyse de peptides présents en mélange. (COTTRELL J.S., Pept. Res., 1997, 7, 115-124). Les peptides caractérisés en fonction de leur masse sont comparés aux protéines et aux protéines associées selon l'invention.

15 **EXEMPLE 4 : Détection d'anticorps spécifiques anti-domaine *env* de HERV-7q.**

L'identification d'un long cadre de lecture ouvert au sein de la séquence *env* de HERV-7q, a permis de déterminer une séquence protéique déduite SEQ ID NO:22 et 35 et figures 18-20 d'une région dudit gène.

20 Les séquences de protéines déduites des séquences ID NO:22, 35 et des figures 18-20 sont positionnées comme suit par rapport à la figure 1 ou à la séquence ID NO:3 :

SEQ ID NO:22 (cadre de lecture 1) et figure 19 : début de la séquence codante : position 7874, fin de la séquence codante 1<sup>er</sup> codon non-sens (position 9493)

25 SEQ ID NO:35 : début de la séquence codante : position 7874, fin de la séquence codante 1<sup>er</sup> codon non-sens (position 9493) (cadre de lecture 1)

Figure 19 : début de la séquence codante : position 6970, fin de la séquence codante 1<sup>er</sup> codon non-sens (position 9493) (cadre de lecture 1)

30 Figure 20 : début de la séquence codante : position 6971, la fin du cadre de lecture est décalée selon le cas de 1, 2 ou 3 codons

Figure 21 : début de la séquence codante : position 6972, la fin du

cadre de lecture est décalée selon le cas de 1, 2 ou 3 codons

Différents peptides correspondant à tout ou partie des SEQ ID NO:22 (voir SEQ ID NO:23-27 et 35) ont été synthétisés par génie génétique afin de tester leur spécificité antigénique vis-à-vis de séra ou de tissus de patients atteints de SEP, par exemple. Brièvement, tout ou partie de la région env de HERV-7q est sous clonée dans les vecteurs pQE30, 31 et 32. Les vecteurs pQE30, 31 et 32 contiennent en 5' du multi-site de clonage les séquences consensuelles pour la transcription (le promoteur fort du bactériophage T5, 2 opérateurs de l'opéron lactose), la traduction (un site d'accrochage ribosomal synthétique). De même, pQE30, 31 et 32 possèdent en 3', le terminateur de transcription du phage  $\lambda$  ainsi qu'un codon "Stop" pour la traduction. L'expression de la protéine s'effectue après transformation dans *E. coli* M15. Le plasmide pQE30, 31 et 32 possèdent en amont du site de polyclonage la séquence codante pour une suite de 6 histidines présentant une affinité pour les ions nickel. Cet enchaînement permet la purification de la protéine chimérique exprimée, par adsorption sur une résine constituée d'un ligand chélatant, l'acide nitrilotriacétique (NTA), chargé de 4 ions nickel (résine NI-NTA, Qiagen).

La transformation s'effectue par électroporation ou traitement au chlorure de calcium. Par exemple, une colonie d'*E. coli* M15 est incubée dans 100 ml de milieu LB contenant 250  $\mu$ g de kanamycine, sous agitation à 37°C jusqu'à l'obtention d'une  $DO_{600}$  de 0,5. Après une centrifugation de 5 minutes à 2000g à 4°C, le culot bactérien est repris dans 30 ml de solution TFB1 (100 mM de chlorure de rubidium, 50 mM de chlorure de manganèse, 30 mM d'acétate de potassium, 10 mM  $CaCl_2$ , 15% glycérol, pH 5.8), à 4°C pendant 90 minutes. Après une centrifugation de 5 minutes à 2000g à 4°C, le culot bactérien est repris dans 4 ml de solution TFB2 (10 mM de chlorure de rubidium, 10 mM de MOPS, 75 mM  $CaCl_2$ , 15% de glycérol, pH 8). Les cellules peuvent être gardées à -70°C par aliquot de 500  $\mu$ l. 20  $\mu$ l de la ligation et 125  $\mu$ l de cellules compétentes sont mélangés et placés dans la glace 20 minutes. Après un choc thermique de 42°C pendant 90 secondes, les cellules sont agitées 90 minutes à 37°C dans 500 ml de milieu Psi-broth (milieu LB complété par 4 mM de  $MgSO_4$ , 10mM de chlorure de potassium). Les cellules transformées sont étalées sur des boîtes LB-agar complétées par 25  $\mu$ g/ml de kanamycine, et 100 $\mu$ g/ml

d'ampicilline, et les boîtes sont incubées une nuit à 37°C.

Les clones potentiellement recombinants sont repiqués de manière ordonnée sur un filtre de nylon déposé sur une boîte LB-agar complémentée par 25 µg/ml de kanamycine et 100 µg/ml d'ampicilline. Après une nuit à 37°C, les clones  
5 recombinants sont repérés par hybridation de l'ADN plasmidique avec la sonde nucléotidique amplifiée par PCR avec le couple d'amorces selon SEQ ID NO:45 et SEQ ID NO:46.

Une colonie indépendante, contenant l'insert, est inoculée à 20 ml de milieu LB complémentée par 25 µg/ml de kanamycine et 100 µg/ml d'ampicilline.  
10 Après une nuit à 37°C sous agitation, 500 ml de même milieu sont incubés au 1/50<sup>e</sup> par cette préculture jusqu'à l'obtention d'une D0<sup>600</sup> de 0,8, puis 1 à 2 mM final d'IPTG est ajouté. Après 5 heures, les cellules sont centrifugées 20 minutes à 4000 g.

Une partie du culot cellulaire est repris dans 5 ml de tampon de sonication (50 mM de phosphate de sodium pH 7,8, 300 mM NaCl) puis placé dans la  
15 glace. Après une rapide sonication, les cellules sont centrifugées 20 minutes à 10000 g. Une partie du culot cellulaire est repris dans 10 ml d'une solution 30 mM Tris/HCl-20% sucrose pH8. Les cellules sont incubées 5 à 10 minutes sous agitation, après adjonction de 1 mM EDTA. Après une centrifugation de 10 minutes à 8000 g à 4°C, le culot est repris dans 10 ml de 5 mM de MgSO<sub>4</sub> glacé. Après 10 minutes dans  
20 la glace sous agitation, les cellules sont centrifugées 10 minutes à 8000 g à 4°C.

Le culot est repris par 5 ml/g dans du tampon A (6 M GuHCl (chlorhydrate de guanidine), 0,1M phosphate de sodium, 0,01M Tris/HCl, pH 8), 1 heure à température ambiante. Le lysat est centrifugé 15 minutes à 10000 g à 4°C, et le surnageant est complémenté par 8 ml de résine Ni-NTA, prééquilibrée dans du tampon  
25 A. Après 45 minutes à température ambiante, la résine est coulée dans une colonne, lavée par 10 fois le volume de la colonne par du tampon A puis par 5 fois le volume de la colonne par du tampon B (8 M urée, 0,1 M phosphate de sodium, 0,01 M Tris/HCl, pH 8). La colonne est lavée par du tampon C (8 M urée, 0,1M phosphate de sodium, 0,01 M Tris/HCl, pH 6,3) jusqu'à ce que l'A280 soit inférieur à 0,01. La  
30 protéine recombinante est éluée par 10 à 20 ml de tampon D (8 M urée, 0,1 M phosphate de sodium, 0,01 M Tris/HCl, pH 5,9) puis par 10 à 20 ml de tampon E (8 M

urée, 0,1 M phosphate de sodium, 0,01 M Tris/HCl, pH 4,5), puis par 20 ml de tampon F (6 M HCl, 0,2 M acide acétique). Après une analyse en SDS-PAGE, la ou les fractions purifiées contenant la protéine chimérique ont permis l'obtention d'anticorps chez le lapin. Les anticorps obtenus sont testés par Western-blot après révélation par

5 un anticorps secondaire couplé à la phosphatase alcaline.

Des anticorps sont obtenus de la même manière, à partir de peptides synthétisés chimiquement selon la technique de Merrifield (G. Barany and B. Merrifield, 1980, dans *The peptides*, 2, 1-284, E. Gross et J. Meienhofer, Academic Press, New York).

10 Les anticorps spécifiques obtenus sont utilisés à fin de détection de l'expression sérique ou tissulaire de tout ou partie des séquences rétrovirales endogènes selon l'invention, dans les cas normaux et pathologiques.

Les protéines d'origine sérique ou tissulaire, sont séparées sur gel d'acrylamide-SDS puis transférées sur un filtre de nitrocellulose à l'aide d'un appareil

15 Novablot 2117-2250 (LKB). Le transfert est effectué sur une feuille de Hybond C-extra (Amersham) en utilisant un tampon CAPS 100 mM pH 11, méthanol, eau (V/V/V: 1/1/8) contenant 1 mM de  $\text{CaCl}_2$ . Après un transfert de 1 heure à 0,8 mA/cm<sup>2</sup>, la feuille est saturée une heure à température ambiante dans du PBS-0,5 % gélatine. La feuille est mise en présence de l'anticorps spécifique à la concentration de

20 1/1000 dans du PBS-0,25 % gélatine. Au bout de 2 heures, le filtre est lavé 3 fois 15 minutes dans du PBS-0,1 % de Tween-20, puis le filtre est incubé 30 minutes en présence d'un anticorps secondaire couplé à la phosphatase alcaline (Promega), dilué au 1/7500 dans du PBS-0,25% gélatine. Après trois lavages dans du PBS-0,1 % de Tween-20, le filtre est équilibré dans un tampon (100 mM de Tris-HCl pH 9,5, 100

25 mM de NaCl, 5 mM de  $\text{MgCl}_2$ ). La révélation est effectuée en présence de 45 µl de NBT à 75 mg/ml et 35 µl de BCIP à 50 mg/ml, pour 10 ml de tampon de phosphatase alcaline.

Les protéines chimériques obtenues par génie génétique, sont utilisées aussi à fin de tests d'activité biologique, comme par exemple pour le test

30 d'activité biologique du peptide de type CKS-17 identifié dans le domaine *env* de

HERV-7q (figure 5).

**EXEMPLE 5 : Obtention de sondes ribonucléiques codant pour les séquences *env* de HERV-7q.**

Les fragments de PCR obtenus sont sous clonés dans le plasmide  
5 PGEM 4Z (Promega) qui possède de par et d'autre de son site de polyclonage, les séquences promotrices pour les ARN polymérase SP6 et T7.

La méthode de compétence utilisée est l'électroporation. Le plasmide et le fragment de PCR sont hybridés dans un rapport de 50 ng de vecteur (coupé à Sma I) pour 100 ng de fragment de PCR (rendu à bout franc par traitement par le fragment  
10 de Klenow de l'ADN polymérase). L'incubation a lieu une nuit à 22°C, dans le tampon de ligation (66 mM Tris-HCl pH 7,5, 5 mM MgCl<sub>2</sub>, 1 mM dithioerythritol, 1 mM ATP) en présence de 1u. de T4 ADN ligase puis est arrêtée par dénaturation 10 minutes à 65°C. Parallèlement, la souche d'*E. Coli* JM 105 estensemencée une nuit à 37°C dans du milieu LB. Cette préculture est diluée au 1/500 et placée à 37°C jusqu'à  
15 une DO<sup>600</sup> égale à 1. Pour la suite du mode opératoire les cellules seront toujours conservées au froid. Après une centrifugation de 5 minutes à 3500 g à 4°C, le culot cellulaire est resuspendu dans 1/4 vol. d'eau glacée ultra-pure. Cette étape est répétée 5 à 6 fois. Puis le culot est resuspendu dans 1/4000 vol. d'eau; 10 % de glycérol stérile sont ajoutés permettant la conservation des cellules électrocompétentes, par aliquots  
20 de 10 µl à 20°C. A 50 µl de cellules électrocompétentes est ajouté 1 µl de la ligation ; le tout est soumis à une décharge électrique de 12,5 kV/cm, appliquée pendant 5,8 ms. Les cellules sont rapidement remises en suspension dans le milieu SOC, incubées 1 heure à 37°C, puis étalées, en présence de 2% X-Gal dans du diméthylformamide, et 10 mM d'IPTG, sur une boîte de gélose LB-agar supplémentée en ampicilline (100  
25 µg/ml). Après une nuit à 37°C, les clones blancs potentiellement recombinants, sont repiqués de manière ordonnée sur une boîte LB/ampicilline et parallèlement sur un filtre de nylon déposé sur une boîte LB/ampicilline. Ces deux boîtes sont incubées une nuit à 37°C. Les clones recombinants sont alors repérés par hybridation avec une sonde nucléique amplifiée par PCR avec le couple d'amorces selon SEQ ID NO:45 et  
30 SEQ ID NO:46 et marquée à la digoxygénine.

Les clones recombinants sont cultivés dans 50 ml de milieu



LB/ampicilline (100 µg/ml) en agitation pendant une nuit à 37°C. Après une centrifugation à 3500 g pendant 15 minutes à 4°C, le culot bactérien est repris dans 4ml de tampon P1 (50 mM Tris-HCl, 10mM EDTA, 400 µg/ml RNase A, pH 8) et 4ml de tampon P2 (200 mM NaOH, 1% SDS). Le mélange est incubé à température ambiante pendant 5 minutes. Après adjonction de 4ml de tampon P3 (2.55 M d'acétate de potassium, pH 4,8) le mélange est centrifugé à 12000 g pendant 30 minutes à 4°C. Le surnageant est appliqué sur une colonne Qiagen-type 100, prééquilibrée avec 2 ml de tampon QBT (750 mM NaCl, 50 mM MOPS, 15% éthanol, pH 7). la colonne est lavée avec 2 fois 4ml de tampon QC (1M NaCl, 50 mM MOPS, 15 % éthanol, pH 7) et l'ADN est élué avec 2ml de tampon QF (1,2 M NaCl, 50mM MOPS, 15 % éthanol, pH 8). L'ADN est précipité avec 0,8 vol. d'isopropanol, et centrifugé à 12000 g à 4°C pendant 30 minutes. Le culot est lavé avec de l'éthanol à 70 % glacé. puis l'ADN plasmidique est repris par 2 fois 150 µl de tampon TE.

Les sondes ribonucléiques sont utilisées comme sondes spécifiques, en particulier pour la détection des transcrits exprimés par les séquences rétrovirales endogènes selon l'invention.

**EXEMPLE 6 : Construction d'une souris transgénique contenant tout ou partie du gène de l'envérine.**

Une souris transgénique contenant tout ou partie de la séquence HERV-7q (SEQ ID NO:3) est construite afin d'identifier les séquences responsables de la spécificité tissulaire, et pour évaluer le rôle de tout ou partie des motifs rétroviraux endogènes de type HERV-7q, en particulier tout ou partie des motifs peptidiques de l'envérine. La technique de micro-injection utilisée se réfère à la technique classique (Hogan et coll., (1994), Manipulating the mouse embryo, Cold Spring Harbor, Cold Spring Harbor Laboratory Press) ou à ses équivalents. Des formes identiques à la molécule humaine normale de motifs de type HERV-7q, dont l'envérine, ou des formes mutées, délétées, présentant des insertions ou tronquées sont testées afin de déterminer les motifs critiques tant sur le plan normal que pathologique, et plus particulièrement au cours du développement foetal et lors des processus tumoraux.

**30 Bibliographie :**

- Benit L. et al., 1997. Cloning of a new murine endogenous retrovirus MuERV-L, with

- strong similarity of the human HERV-L element and with a *gag* coding sequence closely related to the Fv1 restriction gene. *J. Virol.* 71, 5652-5657.
- Coffin J.M. 1985. Endogenous retrovirus. In: "RNA tumor viruses" (Weiss R.A., Varmus H.E., Teich N.M., and Coffin J.M. eds), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Conrad B., Weissmahr R.N., Boni J., Arcari R., Schupbach J., and Mach B. 1997. A human endogenous retroviral superantigen as candidate autoimmunogene in type 1 diabetes. *Cell* 90, 303-313.
- Covey S.N. 1986. Amino acid sequence homology in *gag* region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus. *Nucleic Acids Res.* 14, 623-633.
- Hertig C., Coupar B.E., Gould A.R., and Boyle D.B. 1997. Field and vaccine strains of fowlpox virus carry integrated sequences from the avian retrovirus, reticuloendotheliosis virus. *Virology* 235, 367-376.
- Hohenadl C., Leib-Mösch C., Hehleemann R., and Erfle Y. 1996. Biological significance of human endogenous retroviral sequences. *J. Acqui. Imm. Def. Synd. Hum. Retrovir.* 13, S268-S273.
- Kulkoski J.K., Jones S., Katz R.A., Mack J.P.G., and Skalka A.M. 1992. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. *Mol. Cell. Biol.* 12, 2331-2338.
- La Mantia G. et al, N.A.R., 1991, 19, 7, 1513-1520
- Patience C., Wilkinson D.A., and Weiss R.A. 1997. Our retroviral heritage. *Trends Genet.* 13, 116-120.
- Pearson W.R. 1994. Using the FASTA program to search protein and DNA sequence databases. *Methods Mol. Biol.* 24, 307-331.
- Perron H., Garson J.A., Bedin F., Beseme F., Paranhos-Baccala G., Komurian-Pradel F., Mallet F., Tuke P.W., Voisset C., Blond J.L., Lalande B., Seigneurin J.M., Mandrand B. and the Collaborative Research Group on Multiple Sclerosis. 1997. Molecular identification of a novel retrovirus repeatedly isolated from patients with multiple sclerosis. *Proc. Natl. Acad. Sci. USA* 94, 7583-7588.

- Tönjes R.R. et al., J. AIDS and Hum. Retrovirol, 1996, 13, S261-S267
- Vitelli R., Chiarillo M., Lattero D., Bruni C.B., and Bucci C. 1996. Molecular cloning and expression analysis of the human Rab7 GTP-ase complementary deoxyribonucleic acid. Biochem. Biophys. Res. Commun. 229, 887-890.
- 5 - Weber L.T., Miller M., Jaskolski M., Leis J., Skalka M., and Wlodawer A. 1989. Molecular modeling of the HIV-1 protease and its substrate binding site. Science 243, 928-931.
- Wilkinson D., Mager D.L., and Leong J.A.C. 1994. Endogenous human retroviruses. In: "The Retroviridae" (Levy, J.A. ed), Plenum Press New York, , Vol. 3, 465-535.
- 10 - Xiong Y., and Eickbush, T. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. 9, 3353-3362.

Ainsi que cela ressort de ce qui précède, l'invention ne se limite nullement à ceux de ses modes de mise en œuvre, de réalisation et d'application qui viennent d'être décrits de façon plus explicite ; elle en embrasse au contraire toutes les  
15 variantes qui peuvent venir à l'esprit du technicien en la matière, sans s'écarter du cadre, ni de la portée, de la présente invention.

### REVENDICATIONS

1°) Fragment d'acide nucléique purifié, caractérisé en ce qu'il comprend tout ou partie d'une séquence codant pour une séquence rétrovirale endogène humaine, qui présente au moins des motifs rétroviraux de type *env*, répondant à  
5 la séquence SEQ ID NO:1 ou à une séquence présentant un niveau d'homologie avec ladite séquence SEQ ID NO:1 supérieur ou égal à 80% sur plus de 190 nucléotides ou supérieur ou égal à 70% sur plus de 600 nucléotides pour les domaines de type *env*.

2°) Fragment d'acide nucléique selon la revendication 1, caractérisé en ce qu'il présente à la fois des motifs rétroviraux correspondant à un domaine *env* et  
10 répondant à la séquence SEQ ID NO:1 et des motifs rétroviraux correspondant à un domaine *gag* et répondant à la séquence SEQ ID NO:2 ou à une séquence présentant un niveau d'homologie supérieur ou égal à 80% sur plus de 190 nucléotides ou supérieur ou égal à 70% sur plus de 600 nucléotides pour les domaines de type *env* et un  
15 ou égal à 70% sur plus de 1200 nucléotides pour les domaines de type *gag*, lesquels motifs ne présentent aucune insertion ou délétion supérieure à 200 nucléotides.

3°) Fragment d'acide nucléique, caractérisé en ce qu'il comprend un segment d'une séquence selon la revendication 1 ou la revendication 2 et notamment les séquences SEQ ID NO:3-22, 28 et 61, les séquences nucléiques complémentaires et  
20 les séquences inverses complémentaires des séquences précédentes ainsi que les fragments issus des régions codantes des séquences précédentes correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires.

4°) Transcrits, caractérisés en ce qu'ils sont générés à partir des séquences selon l'une quelconque des revendications 1 à 3.

25 5°) Réactif de diagnostic pour la détection différentielle de séquences nucléiques endogènes humaines complètes ou partielles, présentant des motifs rétroviraux, sélectionnés parmi les séquences SEQ ID NO:1 et/ou SEQ ID NO:2, caractérisé en ce qu'il est sélectionné dans le groupe constitué par les séquences SEQ ID NO:1-22, 28, 37-57, 59-61 et 121-122, les séquences nucléiques complémentaires et les séquences inverses complémentaires des séquences précédentes, par  
30 les fragments nucléotidiques capables de définir ou d'identifier les séquences SEQ ID

NO:1 et/ou SEQ ID NO:2 et toute séquence flanquante ou les chevauchant ainsi que par les fragments issus des régions codantes des séquences SEQ ID NO:1-22 et 61, correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires, éventuellement marquées avec un marqueur approprié.

- 5                   6°) Réactif selon la revendication 5, caractérisé en ce qu'il est choisi dans les régions situées entre les nucléotides 3065 et 4390, les nucléotides 6965 et 9550 ou les nucléotides 2502-2865 de la SEQ ID NO:3.

- 7°) Réactif selon la revendication 5, caractérisé en ce qu'il est sélectionné parmi les séquences SEQ ID NO:37-57, 59-60 et 121-122 et en ce qu'il est apte  
10 à être utilisé comme amorce.

                  8°) Réactif selon la revendication 5, caractérisé en ce qu'il est sélectionné parmi les séquences suivantes :

- un fragment de 1505 nt amplifié par le couple d'amorces SEQ ID NO:37 et SEQ ID NO:38 (amorces G1F et G1R),
  - 15               - un fragment de 2529 nt amplifié par le couple d'amorces SEQ ID NO:45 et SEQ ID NO:46 (amorces E1F et E1R)
  - un fragment de 182 nucléotides répété deux fois, situé en amont du domaine *gag* aux positions 2502-2611/2613-2865
- et en ce qu'il est apte à être utilisé comme sonde.

- 20               9°) Réactif selon la revendication 5, caractérisé en ce qu'il est choisi dans le groupe constitué par les fragments codants ou non-codants pour tout ou partie de l'envérine et notamment les fragments comprenant au moins 14 nucléotides et plus particulièrement les fragments codant pour la partie C-terminale de l'envérine, soit à partir de l'acide aminé 291, soit à partir de l'acide aminé 321, à compter du codon  
25 codant pour la première méthionine.

                  10°) Procédé de détection rapide et différentiel des séquences nucléiques rétrovirales endogènes de type *env* ou *env* et *gag*, de leurs variants normaux ou pathologiques, par hybridation et/ou amplification génique, réalisé à partir d'un échantillon biologique, lequel procédé est caractérisé en ce qu'il comprend :

- 30               (a) une étape dans laquelle l'on met en contact un échantillon biologique à analyser avec au moins une sonde selon la revendication 5, la revendication 6

ou la revendication 8 et

(b) une étape dans laquelle on détecte par tout moyen approprié le ou les produits résultants de l'interaction séquence nucléotidique-sonde.

11°) Procédé de détection selon la revendication 10, caractérisé en ce qu'il comprend :

\* préalablement à l'étape (a) :

. une étape de préparation du tissu ou du liquide biologique concerné,

. une étape d'extraction de l'acide nucléique à détecter, et

10 . au moins un cycle d'amplification génique mis en œuvre à l'aide d'au moins un réactif selon l'une quelconque des revendications 5 à 7 et

\* postérieurement à l'étape (b) :

15 . une étape de comparaison des séquences nucléiques obtenues dans ledit échantillon biologique avec les séquences rétrovirales endogènes humaines selon l'une quelconque des revendications 1 à 3, par tout moyen approprié et notamment par séquençage, Southern-blot, coupure de restriction, SSCP ou toute autre méthode permettant d'identifier une insertion ou une délétion ou encore une simple mutation entre les différentes séquences comparées.

12°) Procédé de détection des transcrits selon la revendication 4, caractérisé en ce qu'il comprend :

- le prélèvement des ARN messagers provenant d'échantillons biologiques témoins et d'échantillons analogues prélevés chez des patients et

20 - l'analyse qualitative et/ou quantitative desdits ARNm, par hybridation *in situ*, par dot-blot, Northern-blot, RNase mapping ou RT-PCR, à l'aide d'un réactif de diagnostic selon l'une quelconque des revendications 5 à 9.

13°) Séquences chimères, caractérisées en ce qu'elles sont constituées par un fragment de 17 à 40 nucléotides d'une séquence flanquante sélectionnée dans le groupe constitué par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou 30 l'altération est associée à l'expression normale ou pathologique ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences

correspondant à des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite famille HERV-7q dont l'une des extrémités ne peut se trouver à une distance excédant 120 kb, associée à un motif rétroviral endogène de type HERV-7q comprenant entre  
5 17 et 40 nucléotides selon les revendications 1 à 4.

14°) Méthode de détection et/ou d'évaluation d'une sur-expression/sous-expression ou d'une modification d'au moins l'une des séquences ou fragments de séquences rétrovirales endogènes de type HERV-7q et/ou de leurs séquences flanquantes associées, selon l'une quelconque des revendications 1 à 9, caractérisée en  
10 ce qu'elle comprend :

- le dépôt sur un support approprié, de l'ADNc provenant de clones, de produits de PCR obtenus à partir d'ADN génomique, de produits de RT-PCR provenant de transcrits ou encore de séquences oligonucléotidiques spécifiques, lesdites séquences d'ADN étant des séquences ou des fragments de séquences rétro-  
15 virales endogènes de type HERV-7q et/ou leurs séquences flanquantes, constituées par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou l'altération est associée à l'expression normale ou pathologique ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences correspondant à des séquences  
20 nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite famille HERV-7q dont l'une des extrémités ne peut se trouver à une distance excédant 120 kb et/ou une séquence chimère selon la revendication 13,

- l'hybridation dudit support avec au moins une sonde marquée de  
25 manière adéquate obtenue, par exemple, par rétrotransposition d'un mélange d'ARN provenant de cellules, de tissus ou de liquides biologiques provenant de témoins réputés normaux, de membres de populations ethniques différentes, de patients atteints de pathologies souvent associées à une expression de rétrovirus, comme les processus tumoraux, ou comme les maladies auto-immunes, et

30 - la détection des hybrides formés.

15°) Méthode selon la revendication 14, caractérisée en ce que ledit

transcrit ou ADNc est sélectionné dans le groupe constitué par les séquences SEQ ID NO:62-67 et 119 et leurs fragments correspondant à un cadre glissant supérieur ou égal à 14 nucléotides ou leurs séquences complémentaires.

16°) Méthode selon la revendication 14 ou la revendication 15.  
5 caractérisée en ce que ledit support comprend en outre toute séquence rétrovirale endogène ou exogène.

17°) Kit de détection et/ou d'évaluation d'une maladie auto-immune et notamment des neuropathologies à étiologie auto-immune. caractérisé en ce qu'il comprend outre les tampons nécessaires à la mise en œuvre d'un procédé selon l'une  
10 quelconque des revendications 14 à 16 :

- des réactifs A de diagnostic selon l'une quelconque des revendications 5 à 9, et
- des réactifs B constitués par les transcrits et ADNc des séquences génomiques, qui codent pour tout ou partie d'un facteur, dont la fonction, la régulation/dérégulation ou l'altération est associée à l'expression normale ou pathologique  
15 tion/dérégulation ou l'altération est associée à l'expression normale ou pathologique ou à la régulation/dérégulation de motifs appartenant à ladite famille HERV-7q, ces séquences correspondant à des séquences nucléotidiques codant pour des gènes situés dans des régions flanquantes situées en amont et/ou en aval d'une séquence rétrovirale de ladite famille HERV-7q dont l'une des extrémités ne peut se trouver à une distance  
20 excédant 120 kb,

lesquels réactifs sont de préférence fixés sur un support approprié.

18°) Kit selon la revendication 17, caractérisé en ce que lesdits réactifs B sont sélectionnés dans le groupe constitué par les séquences SEQ ID NO:62-67 et 119 et leurs fragments correspondant à un cadre glissant supérieur ou  
25 égal à 14 nucléotides ou leurs séquences complémentaires.

19°) Produits de traduction, caractérisés en ce qu'ils sont codés par une séquence nucléotidique selon l'une quelconque des revendications 1 à 4.

20°) Peptide, caractérisé en ce qu'il est susceptible d'être exprimé à l'aide d'une séquence nucléotidique sélectionnée dans le groupe constitué par les  
30 séquences SEQ ID NO:1-22, 28 et 61 selon l'une quelconque des revendications 1 à 4.

21°) Peptide selon la revendication 20, caractérisé en ce qu'il



englobe les peptides dérivés comprenant entre 5 et 540 aminoacides et notamment un fragment de 538 aminoacides, commençant à la première méthionine de la séquence SEQ ID NO:26 (envérine).

- 22°) Peptide selon la revendication 20 ou la revendication 21,
- 5 caractérisé en ce qu'il est sélectionné dans le groupe constitué par :
- . les séquences SEQ ID NO:23-36 ;
  - . la séquence SEQ ID NO:58 ;
  - . un fragment C-terminal de la séquence SEQ ID NO:26, soit à partir de l'acide aminé 291, soit à partir de l'acide aminé 321, à compter de la première
- 10 méthionine de la séquence SEQ ID NO:26 ;
- un peptide de type CKS-17/CKS-25 présent dans l'une des séquences SEQ ID NO:23-36 ou 58 ; et
  - les peptides présentant une affinité avec un des haplotypes du système HLA de classe I ou de classe II et notamment les fragments 399-471, 244-271
- 15 de l'envérine, ainsi que les peptides de séquence SEQ ID NO:68-118, conformément au Tableau I.

- 23°) Peptide selon l'une quelconque des revendications 20 à 22, caractérisé en ce qu'il est obtenu à partir des séquences nucléiques selon l'une quelconque des revendications 1 à 4, dans lesquelles au moins un codon non-sens peut
- 20 être remplacé par un codon codant pour l'un des aminoacides suivants : Phe (F), Leu (L), Ser (S), Tyr (Y), Cys (C), Trp (W), Gln (Q), Arg (R), Lys (K), Glu (E) ou Gly (G).

- 24°) Compositions immunogènes ou vaccinales, pour la protection contre les maladies auto-immunes, notamment chez les sujets à risque, caractérisée en
- 25 ce qu'elle comprend au moins un peptide comprenant au moins un motif de type CKS et/ou au moins un motif sélectionné dans le groupe constitué par les peptides présentant une affinité avec un des haplotypes du système HLA de classe I ou de classe II et au moins un véhicule pharmaceutiquement acceptable.

- 25°) Composition selon la revendication 24, caractérisée en ce que
- 30 ledit peptide présentant une affinité avec un des haplotypes du système HLA de classe I ou de classe II, est sélectionnée dans le groupe constitué par les peptides tels que

définis dans le Tableau I.

26°) Composition selon la revendication 24 ou la revendication 25, caractérisée en ce que ledit peptide présente la séquence SEQ ID NO:120.

27°) Anticorps, caractérisé en ce qu'il est dirigé contre l'un ou  
5 plusieurs des peptides selon l'une quelconque des revendications 20 à 23.

28°) Composition pharmaceutique, caractérisée en ce qu'elle comprend des anticorps neutralisants produits à partir des peptides du Tableau I (SEQ ID NO:68-118) et leurs homologues.

29°) Procédé de dépistage immunologique différentiel de séquences  
10 rétrovirales endogènes humaines de la famille HERV-7q normales ou pathologiques, caractérisé en ce qu'il comprend la mise en contact d'un échantillon biologique avec un anticorps selon la revendication 27, la lecture du résultat étant révélée par un moyen approprié, notamment EIA, ELISA, RIA, fluorescence.

30°) Procédé d'identification et de détection de motifs rétroviraux  
15 endogènes, anormalement exprimés dans le cadre de pathologies associées au cancer, ou de neuropathologies en particulier autoimmunes, au premier rang desquelles la sclérose en plaques, caractérisé en ce qu'il comprend l'analyse comparée des séquences extraites d'un échantillon biologique avec les séquences selon l'une quelconque des revendications 19 à 23.

20 31°) Application des séquences selon l'une quelconque des revendications 1 à 9, 13, 14 ou 19 à 23 au diagnostic, au pronostic, à l'évaluation de la susceptibilité génétique, à toutes maladies humaines induites, innées ou acquises en particulier celles à composantes cancéreuses, autoimmunes et/ou à incidence neurologique, comme la sclérose en plaques, les syndromes associés et les maladies neuro-  
25 dégénératives où intervient tout ou partie des séquences selon l'une quelconque des revendications 1 à 5 et des formes endogènes ou exogènes apparentées.

32°) Séquences nucléiques hybrides, caractérisées en ce qu'elles comprennent des séquences ou motifs selon l'une quelconque des revendications 1 à 9, combinés avec des séquences ou motifs d'origine endogène ou d'origine ou induits  
30 de manière exogène.

33°) Vecteur recombinant de clonage ou d'expression, caractérisé en

ce qu'il comprend une séquence nucléique selon l'une quelconque des revendications 1 à 4.

34°) Composition immunogène ou vaccinale, caractérisée en ce qu'elle comprend un vecteur incluant au moins une séquence nucléique codant un peptide tel que défini dans le Tableau I, éventuellement associée à une séquence codant un motif de type CKS-17.

35°) Vecteur de thérapie génique, caractérisé en ce qu'il comprend tout ou partie des séquences nucléiques rétrovirales endogènes de type HERV-7q selon l'une quelconque des revendications 1 à 4.

10 36°) Vecteur selon la revendication 35, caractérisé en ce que lesdites séquences sont sélectionnées dans le groupe constitué par les séquences SEQ ID NO:2, 20 et 21.

37°) Animaux transgéniques, caractérisés en ce qu'ils comprennent tout ou partie d'une séquence de type HERV-7q (SEQ ID NO:1-22, 28 et 61).

1/64

CCTTGGGGGGGGCTTCTTCTGGGATGAGGGCAAACGCTGGAGATACAGCAATTATCTTGCACCTGAG  
 AGACAGGACTAGCTGGATTTCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACCACGCTCCAC  
 CTTTAAACACGGGGCTTGCACCTTAGCTCACACCTGACCAATCAGAGAGCTCCTAAATGCTAATTAGGCA  
 AAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGGCTGAGAGCACAGCAGGAGGACAAATCGGGATA  
 TAAACCCAGGCATTGAGCTGGCAACAGCAGCCCCCTTTGGGCTCTTCCCTTTGTATGGGAGCTGTTTC  
 ATGCTATTCTACTCTATTAAATCTTGAACCTGCACCTCTTCTGGTCCATGTTTCTTACGGCTCGAGCTGAGCT  
 TTTGCTCACCGTCCACCACTGCTGTTTGGCCACCACCGCAGACCTGCCGCTGACTCCCATCCCTCTGGATCCT  
 GCAGGGTGTCCGCTGTGCTCTGATCCAGCAGGGCGCCCAATTGCCGCTCCCAATTGGGCTAAAGGCTTGCCT  
 TTGTTCTGCACGCTTAAGTGCTGGGTTTGTCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTC  
 TCTTCTGTGACCCAGCGCTTCTAATAGAACTATAACACTTACCACATGGCCCAAGATTCCATTCTTGGAA  
 CCGTGAGGCCAAGAACTCCAGCTCAGAGAAATACGAGGCTTGGCCACCATCTTGGAAAGCGGCTGCTACCATCT  
 TGAAGTGGTTCACCACTCTTGGGAGCTCTGTGAGCAAGGACCCCGGTAAACATTGGCAACCAAGAA  
 CGGACATCCAAAGTGGTGAGTAATATTGGACCACTTTCATTGCTATTCTGTCTTATCTCTTAGAATTC  
 GAGGAAAAATACGGGCACCTTGTGGCCAGTAAACACGATTAGTGTGGCCACCGGACTTAAGACTCAGGTGT  
 GAGGCTATCTGGGAAGGGCTTTCTAACACCCCAACCTTCTGGGTTGGGACTTGGTTTGCCTCAAGCC  
 AGCTTCCACTTTCAGTTTCTTGGGAAGCCGAGGGCCGACTAGAGGAGAGAGCTGTCTCTGAACTCCC  
 GGCAGTAGCCGCTTGAATCATGGTGTAGCCAGAACTCTCAACAGTCCGCTATGCATGACCCCTATCTTTC  
 CTTCTGACCCATACCTCTGGGTCCCAACCACTTCTTCAAAGTGTAGCCCAAAATTCCTCTTACCTC  
 TGAATATACTTCTCTGATCCCTGCTCTAGGTACTATTGGTTGACACTTCCATTTCTCTAGCAAGTGT  
 ATCTCCAAAGGATCTAAGGAAGCTCTGGCTGCTCTTGGCAGCTTAGGCTATAACCCAGGAGCTTAT  
 CCTGGTGTCCCTCCCAATTTAGGCATACAGCTTTCAGATGGGCACTTATGAGGACCCACTCCCCACAC  
 CCTTGCAGGGGCCCCAAGTTTGAATGGCTGAGGGAAGAGAGACAGAGGAGAGAGAGAAATGGAGGA  
 AGAAG  
 AGAGAGAGTCAAAGAGAGAAAG  
 GGGTAAATTTAAACCTGTACTTGATAATGAAGGTCTCTGTGACCTTATAGCACTCCAATCCACTTTG  
 TGGTCAGTGAATAAGAGCATAGGCCGAAGCACTGAGGCCATTGACACCCGTAAGTCTTCTTAAAGTGA  
 TCCCTTAACCCAGTAACCCGAGATGGACCAATGCACTGAGTGGTGAAGCACTGCTTTGCTAAAGTAGA  
 AAAAGTAACTTTAGAGGAACCTCATTTGTGAGCACCTTCACTGTTGAGAAATATTCTAATAAAAAAGCA  
 AACAGGTAGCTTACTAATCAAAATCTTAAAGTATGGGGCTATTCTGTAGAAAAAGTAAATGTAATCCA  
 ACCACTGATAATTCCTTAAACCCAGCAGATTCTTAAACCCGATTTAAATCTTAAATACCATCAAGGTCG  
 ACCAGACTAGGCGGAATTCCTTCAAGGAGAGAGATAGATGGTTCCTCCAGGTGATTGAGGAAAAAAC  
 CACAATGGGTATTAGTAATGATACGGGACTCTTGTGGAAGCAGAGTTAGAAAAATTGCCTAATAACTGG  
 TCTCTCAACGCTGTGAGCTTTTGCATCAGCCAAAGCTTAAAGTACTTACAGAACTCAAAAGACTATCTCA  
 ATCTGATTCAAAAGGTTAGTACACCTCTCTGTATGCACTTGCATAGAACTTGTATTATGGGAATGCAT  
 CTTGATGGGGAGCTGGGTTGTTAATAAGTAAAGGACCACTAGAACTCCAGCAGCCAGACCCCTTCTTTTGGTCAAGAAA  
 GCAATGTTGGGATGCTGGTAAAGGACCACTAGAACTCCAGCAGCCAGACCCCTTCTTTTGGTCAAGAAA  
 GGGGGGAAAGGGGTGACGAGCTGCTACCTGGTAAGCATAACTTAATCCGATAAACAGAGGTCCATGGGTGG  
 TTACGCACTGCAAGGAACTTCAACCTGAGCACAAGGCAATGTTGGGACGCTGGTAAAGGACCACTAG  
 AATCCAGCAGCTGGACCCCTTCTTTTGGTCAAGAGGCAAGGAAACAGGTGCAAGACTGCAACATCAG  
 TGAGCATTAATTAATTCGATAGCAGAGGTCCATGGGTGGTATGCACTTGGAAAGAAATAGCAATAGGACC  
 ATAGAGGACACTCCAGGACTAAAGCTCATCGGAAATGACTAGGGTGTGCTGGCATCCCTATGTTCTTTTTC  
 AGATGGGAAAGCTTCCCCGAAAGACAAACGCCCCCTAAGACGATTCTGGAGAAATGGGACCAATTTGACC  
 CTCAGACACTAAGAAAGAAACGACTTATTTCTTCTGCACTGCCCCCTGGCACTCTGAGGAAAGTAAAT  
 TATAACCACTCTTACAGCTAGACCTCTTTTGTAGAAAAGGCAATGGAGTGAAGTCCATAAGTACAACCT  
 TTCTTTTCAATTAAGAGACAACCTACAATATGTAAGAAAGTGTGATTATGCCCTACAGGAAGCTTACAGAGT  
 CTACCTCTCTATCCAGCATCCCCGACTCTTCCCCAACTAATAAGGACCCCTTCAACCCAAATGGTCCA  
 AAGGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTGGCAATATTCCCCAATTATGACCCCTCCAAGC  
 AGTGGGAGGAAGAAATTGGGCCAGCAGAGTGCATGTGCTTTTCTCTCCAGACTTAAAGCAAAATAA  
 AACAGACTTAGGTAATTTCTAGATAACCTGTATGGCTATATTGATGTTTCAAGGGTTAGGCAATTTCTT  
 TGATCTGACATGGAGAGATATAATGTCACTGCTAAATCAGACACTAACCCCAATGAGAGAAAGTGGCACC  
 AACTGACGCTGAGAGTTTGGCGATCTCTGTATCTCAGTCAGGTCAATGATAGGATGACAACAGAGGAAAG  
 AGAATGATTTCCCAAGGCCAGCAGGAGTTCCAGTCTAGACCTCAATTGGGACACAGAAATCAGAACTGG  
 AGATTGGTGTGACAGACTTGTCTAATTTGTGTGCTAGAAGGACTAAGGAAACTAGGAAGAGTCTATGAA  
 TTACTCAATGATGTCACCATAAACAGGGAAGGGAAGAAAATCCTACTGCTTTCTGGAGAGACTAAGGGA  
 GGCAATTGAGGAAGCTGCTCTCTGTCACTGACTTCTTGAAGGCCAACTAATCTTAAAGCGTAAGTTAT  
 CACTCAGTCAGTGCAGACATTAGAAAAAACTTCAAAAGTCTGCCGTAGGCCCGGAGCAAACTTAGAAAC  
 CCTATTGAACCTTGGCAACCTCGGTTTTTATAATAGAGATCAGGAGGAGCAGCGGAAACAGGACAAACGGGA  
 TTAATAAAAGGCCACCGCTTTAGTATGACCTCAGGCAAGTGGACTTTGGAGGCTCTGGAAAAGGGA  
 GCTGGGCAAAATGAATGCTTAATAGGCTTGTCTCCAGTGGGCTTACAAGGACACTTAAATAAGATTGTCT  
 CAAGTAGAAGTAAGCCGCCCCCTCTGCTATGCCCCCTATTTCAGGGAATCACTGGAAGGCCACTGCCCCA  
 GGGGACAAAGGCTCTGAGTCAAGGCACTAACCCAGATGATCCAGCAGCAGGACTGAGGGTGGCTGGGGC  
 AAGCGCCATCCATGCCATCACCCTCAGAGGCCCCGGGTATGCTTGACATTGAGGGCCAGGAGGTTGTCT  
 CTTGGACACTGGTGGGCTCTTCTAGTCTTACTCTTCTGCCCCGACAAGTCTCTCCAGATCTGTCACTAT  
 CTGAGGGGCTCCTAAGACGGGCTGCTAGTACTTCTCTCCAGCCACTAAGTTATGACTGGGAGCTTTAT  
 TCTTTTCACTGCTTTCTAATTTATGCTTGAAGGCCCACTACCTTGTAGGAGAGACATTCTAGCAAAAG  
 CAGGGGCTTATACACCTGAACATAGGAGAAAGGAAACACCCGTTTGTGTCCTCTGCTTGAAGGAATTA  
 ATCTGAAGTCTGGGCAACAGAAAGGACAAATAGGACGCAAGAAATGCCGCTCTGCTTCAAGTTAACTAA  
 AGGATTCACCTCTCTTCCCTACCAAGGCACTACCCCTCAGACCCAAAGGCCCAACAGGACTCCAAAGGA  
 TTGTTAAGGACCTTAAAGCCCAAGGCTAGTAAACCTGAGTAACCCCTGAGTAACCTTCAATTTAGGAG  
 TACAGAAACCAACAGACAGTGGAGGTTAGTGAAGATCTCAGGATTAATCAATGAGGCTGTTGTTCTCTCT  
 AGCCAGCTGTACCTAGCCCTTATCTGCTTTCCCAATACAGAGGAAGCAGAGTGGTTTACAGTCTCTG  
 ACCTCAGGATGCTTCTTCTGATCCCTGTACATCCTGACTCTCAATTTGTTTGCCTTTGAAGATACTT

71  
 143  
 215  
 227  
 359  
 431  
 503  
 575  
 647  
 719  
 791  
 863  
 935  
 1007  
 1079  
 1151  
 1223  
 1295  
 1367  
 1439  
 1511  
 1583  
 1655  
 1727  
 1799  
 1871  
 1943  
 2015  
 2087  
 2159  
 2231  
 2303  
 2375  
 2447  
 2519  
 2591  
 2663  
 2735  
 2807  
 2879  
 2951  
 3023  
 3095  
 3167  
 3239  
 3311  
 3383  
 3455  
 3527  
 3599  
 3671  
 3743  
 3815  
 3887  
 3959  
 4031  
 4103  
 4175  
 4247  
 4319  
 4391  
 4463  
 4535  
 4607  
 4679  
 4751  
 4823  
 4895  
 4967  
 5039  
 5111  
 5183

région  
 répétée  
 R1

régions  
 répétées  
 en tandem  
 R2

domaine  
 gag

domaine  
 pol

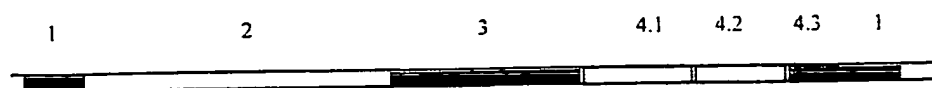
FIGURE 1.1

2/64

CAAAACCAACATCTCAACTCACCTGGACTATTTACCCCAAGGGTTCAGGGATAGTCCCATCTATTTGGCC	5255
AGGCATTAGCCCAAGACTTGAGCAATCTCATACCTGGACACTTGTCTTCGGTAGGTGGATGATTACTT	5327
TTGGCCGCCCATTCAGAAACCTTGTGCCATCAAGCCACCCAGCGCTCTTCAATTTCTCGCTACCTGTGGC	5399
TACATGGTTTCCAAACCAAAGGCTCAACTCTGCTCACAGCAGGTTACTTAGGGCTAAAATTATCCAAAGGCA	5471
CCAGGGCCCTCAGTGAGGAACACATCCAGCCTATACTGGCTTATCCTCATCCAAACCTAAAGCAACTAA	5543
GGGGATTCTTGGCGTAATAGGTTTCTGCCGAAAATGGATTCCCAGGTATGGCGAAATAGCCAGGTCATTAA	5615
ATACACTAATTAAAGGAACTCAGAAAGCCAATACCCATTAGTAGAGTGGCAACTGAAGTAGAAGTGGCTT	5687
TCAGGGCCCTAACCCAGCCCAAGCTGTAAAGTTTGCCAAACAGGGCAAGACTTTTCTCATATGTCACAGAAA	5759
AAACAGGAATAGCTCTAGGAGTCTTACACAGATCCGAGGGATGAGCTTGCACCTGTGGCATACCTGACTA	5831
AGGAAATTGATGATAGTGGCAAGGGTTGACCTCATTGTTACGGGTAGTGGTGGCAGTAGCAGTCTTAGTAT	5903
CTGAAGCAGTTAAAATAATACAGGGAAGAGATCTTACTGTGTGGACATCTCATGATGTGAATGGCATACTCA	5975
CTGCTAAAGGAGACTTGTGGCTGTACAGCAACTGTTACTTAAATGTGAGGCTCTATTACTTGAAGGGCCAG	6047
TGCTGCGACTGTGCACTTGTGCAACTCTTAACCCAGCCACATTTCTCCAGACAATGAAGAAAAGATAAAAC	6119
ATAACTGTCAACAAGTAATTTCTCAACCTATGCCACTCGAGGGGACCTTTAGAGGTTCTTTGACTGATC	6191
CCGACCTCAACTTGTATCTGATGGAAGTTCTTTGTAGAAAAGGAGCTTCGAAAAGTGGGGTATGCAGTGG	6263
TCAGTGATAATGGAATACTTGAAGTAATCCCTCACTCCAGGAAC TAGTGCTCAGCTAGCAGAACTAATAG	6335
CCCTCACTTGGGCACTAGAATTAGGAGAAGAAAAGGGCAATATATATACAGACTCTAAATATGCTTACC	6407
TAGTCTCCATGCCCATGACGCAATATGGAAGAAAGGGAAATTCCTAATCTTGAGAGAACCTATCAAC	6479
ATCAGGAAGCCATTAGGAAATATTATTGGCTGTACAGAACTAAAGAGGTGGCAGTCTTACACTGCCGGG	6551
GTCAATCAGAAAGGAAAGGAAAGGGAATAGAAGAGAAGTCCCAAGCAGATATTGAAGCCAAAAGAGCTGCAA	6623
GGCAGGACCCCTCAATTAGAAATGCTTATAAAACAACCCCTAGTATAGGGTAATCCCTCCGGGAAACCAAGC	6695
CCGAGTACTCAGCAGGAGAAACAGAAATGGGGAACCTCAGGAGGACAGTTTCTCCCTCGGGACCGCTAGCC	6767
ACTGAAGAAGGAAAATACTTTTGCCTGCAACTATCCAATGGAATTAATTAACCCCTTCATCAAACTTT	6839
CACCTTAGGCACTCGATAGCCCACTCAGATGGCCAAATCATTATTACTGGACCGGCTTTTCAAACTATC	6911
AAGCAGTAGTACAGGGCTGTGAAAGTGTGCCAGAGAAATAATCCCTGCCTTATCGCCCAAGCTCCTTCAGGA	6983
GAACAAGAACAGGCCATTACCTTGGAGAAGACTGGCAACTGATTTACCCACAAGCCCAACCTCAGGGAT	7055
TTTCAGTATCTACTAGTCTGGGTAGATACTTTCACGGGTGGGCAGAGGCTTCCCTGTAGGACAGAAAAGG	7127
CCCAAGAGGTAAATAAGGCACTAGTTTCAATGAATAATTTCCAGATTTCGGACTTCCCGAGGCTTACAGAGTG	7199
ACAAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCAGGCGTTAGGTATACGATATCACTTAC	7271
ACTGCGCTGAAGGCCACAGTCTCAGGGAAGGTGAGAAAATGAATGAACACTCAAAGGACATCTAAAAA	7343
AGCAAAACCCAGGAACCCACCTCACATGGGCTGCTCTGTGCTATAGCCTTAAAAAGAACTGCAACTTTC	7415
CCCAAAAGCAGGAGCTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTG	7487
ACCCAAGACAGCCAACTTAGTTGAGCATCACTCCTTAGCCAAATATCAACAAGTTCTTAAAAACATTACA	7559
AGGAACCTATCCCTGAGAAGAGGGAAGAAAGAACTATCCACCTTGTGACATGGTATTAGTCAAGTCCCTTCC	7631
CTCTAATTTCCCATCCCTAGATACATCTCGGGAAGGACCTACCCAGTCATTTATCTACCCCAACTGCGGT	7703
TAAAGTGGCTGGAGTGGAGTCTTGGATACATCACTTGAGTCAAACTCGGATACTGCCAAAGGAACCTGA	7775
AAATCCAGGAGACAACGCTAGCTATTCTGTGAACCTCTAGAGGATTGGCGCTGCTCTTCAACAACAACC	7847
AGGAGGAAATAGCTTAAATATATAATCCCATGGCCCTCCCTTATCATATTTTCTCTTTACTGTTCTTT	7919
ACCCCTTTTACTCTCAGTGCACCCCTCCATGCCGCTGTATGACCAAGTACCTCCCTTACCAAGAGTTCT	7991
ATGGAGAATGCAGGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAAACCCACCTT	8063
TTGGACAGGAAAAATGATTAACTCTAGTTGTCTGGAGGACTTGGAGTCACTGTCTGTGGACTTACTTCAC	8135
CCAACTGGTATGTCTGATGGGGTGGAGTTCAAGATCAGGCAAGGAGAAAAACATGTAAAAAGAAATATCTC	8207
CCAACCTACCCCGGTACATGGCACTCTAGCCCCACAAAGGACTAGATCTCTCAAACTACATGAACCCCT	8279
CGTACCCATACTCGCTGGTAAGCCTATTAAATACCACCTCACTGGGCTCCATGAGGTCTCGGCCAAAAA	8351
CCCTACTAATCTGTTGATATGCCCTCCCTGAACTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATG	8423
GAACAACCTCAGCAGAGAAATAACACCACTTCCGTTTATAGGACCTCTTGTTCCTTCAATCTGGAATAAC	8495
CCATACCTCAAACTCAGCTGTGTAAATTTAGCAATACTACATACACAACCACTCCCAATGCATCAGGTG	8567
GGTAACTCTCCACACAAATAGTCTGCCCTACCCCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCG	8639
TTGTTGAATTCGCTCTTCAAGTCTATGTGCTTCTCTCATCTTAGTGGCCCTATGACCTCTACACTGA	8711
ACAGAGATTATACAGTTATGTCATATCTAAGCCCGCAACAAAGAGTACCCATCTTCTCTTTGTTATAGG	8783
ATCAGGAGTGTAGTGCAGTACTGGCTTGGCGGTATCACAACTCTACTCAGTCTACTACAACT	8855
ATCTCAAGAACTAAATGGGGCATGGAACGGGTGCGCGACTCCCTGGTCACTTGCAGATCAACTTAACTC	8927
CCTAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAGACTTGTAAACGCTGAAGAGGGGGAACTGTTT	8999
ATTTTAGGGGAAGAAATGCTGTATATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCG	9071
AGATCGAATACAACTAGAGCAGAGGAGCTTCGAAACACTGGACCCCTGGGGCTCTCAGCCAAAGGATGCC	9143
CTGGATCTCCCTCTCTAGGACCTCTAGCAGCTATAATATTGCTACTCTCTTTGGACCTGTATCTTTAA	9215
CCTCCTTGTAACTTTGCTCTTCCAGAAATCGAAGCTGTAAACTACAAATGGAGCCCAAGATGCAGTCCAA	9287
GACTAAGATCTACCCAGACCCCTGGACCGGCTGTAGCCACGATCTGATGTTAATGACATCAAAAGGCAC	9359
CCCTCTGTAGGAAATCTCAGCTGCACAACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT	9431
TCGGCCAACTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGGACTGAGAGACAGGACTAGCTGGAT	9503
TTCCTAGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACCACATCCACCTTTAAACACGGGGCTTG	9575
CAACTTAGCTCACTGACCAATCAGAGAGCTCACTAAATGCTAATAGGCAAGACAGAGGTAAGAGAA	9647
ATAGCCAATCATCTATTGCTGAGAGCACAGCAGGAGGACAATGATCGGGATATAAACCAGTCTTCGAG	9719
CCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTGTATGGGAGCTCTGTTTTCATGCTATTTCACCTAT	9791
TAAATCTTCAACTGCACCTTCTGCTCCATGTTTCTTACGGCTTAGGCTGAGCTTTCGCTCGCCATCCACC	9863
ACTGCTGTTTGGCGCCACCGCAGACCCGCGCTGACTCCATCCCTCTGGATCATGACGGGTGTCGCTGTG	9935
CTCCTGATCCAGCAGGACCCCATTTGCCCTCCCAATCGGGCTTAAAGGCTTGCCATTGTTCTGCTGATGGCTA	10007
AGTGCCTGGGTTCATCTAATTAGCTGAACACTAGTCACTGGGTTCATGTTTCTCTCTGTGACCCACAG	10079
CTTCTTAAGTACTATAACACTCAGCCGATGGCCCAAGGTTCCATTCTTGAATCCATAAGGCCAAGAACCC	10151
CAGGTGAGAGAACACGAGGCTTGCCACCATCTTGGGAGCTCTGTGAGCAAGAGCCCAAGTAAACACAACA	10223
TGAGGGTGCAAAATGCATGGGCCACTAATGGTAGAGCAAGAAAACAGAGGGCCCTGGTTCTCGAAGGCATC	10295
AGTAGCTGAAATGGCTGCCCTGGATGTCCTATTCTAGGTTTCTTCTGCTGAAGCAGATTAAACCCCTTT	10367
GTTCACTTCTCAAGTAGGGCTTCTATTACAGCCCAATCAATCCCAACCCAGATGACAT	10439
	10500

FIGURE 1.2

3/64



**FIGURE 2**

4/64

```

ACTGAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACC
:
:
:
ACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACC
:
:
:
ACGTCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGC
:
:
:
ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAATGC
:
:
:
TAATTAGGCAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACA
:
:
:
TAATTAGGCAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGCACAGCAGGAGGGACA
:
:
:
ACAATCGGGATATAAACCCAGGCATTGAGCTGGCAACAGCAGCCCCCTTTGGGTCCCTTCCCTTTGTA
:
:
:
ATGATCGGGATATAAACCCAAAGTCTTCGAGCCGGCAACGGCAACCCCC-TTTGGGTCCCTTCCCTTTGTA
:
:
:
TGGGAGCT--GTTTCATGCTATTTCACCTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCT
:
:
:
TGGGAGCTCTGTTTCATGCTATTTCACCTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCT
:
:
:
TACGGCTCGAGCTGAGCTTTTGCTCACCCTCCACCCTGCTGTTTGCCACCACCGCAGACCTGCCGCTGA
:
:
:
TACGGCTTGAGCTGAGCTTTGCTCGCCATCCACCCTGCTGTTTGCCGCCACCGCAGACCCGCCGCTGA
:
:
:
CTCCCATCCCTCTGGATCCTGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCGCCCATTGCCGCTCC
:
:
:
CTCCCATCCCTCTGGATCATGCAGGGTGTCCGCTGTGCTCCTGATCCAGCGAGGCAACCATTGCCGCTCC
:
:
:
CAATTGGGCTAAAGGCTTGCCATTGTTCTTCTGACCGCTAAGTGCCTGGGTTTGTCTAATTGAGCTGAAC
:
:
:
CAATCGGGCTAAAGGCTTGCCATTGTTCTTCTGATGGCTAAGTGCCTGGGTTTGTCTAATTGAGCTGAAC
:
:
:
ACTAGTCACTGGGTTCATGGTTCTTCTTGTGACCCACGGCTTCTAATAGAACTATAACACTTACCACA
:
:
:
ACTAGTCACTGGGTTCATGGTTCTTCTTGTGACCCACAGCTTCTAATAGAGCTATAACACTCACCAGCA
:
:
:
TGGCCCAAGATTCCATTCTTGGAAATCCGTGAGGCCAAGAAGTCCAGGTCAGAGAATACGAGGCTTGCCA
:
:
:
TGGCCCAAGGTTCCATTCTTG-AATCCATAAGGCCAAGAAGTCCAGGTCAGAGAATACGAGGCTTGCCA
:
:
:
CCATCTTGGGAGC
:
:
:
CCATCTTGGGAGC

```

**FIGURE 3**

5/64

IPMALPYHIFLFTVLLPSFTLTAPPPCRCMTSSSPYQEFWLRMQRPGNIDAPSYRSLSKG  
 TPTFTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGVTVCWTYFTQTGMSDGG  
 GVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRHTRLVSLFNTTLTGLHEV  
 SAQNPTNCWICLPLNFRPYVSI PVPEQWNNFSTEINTTSVLVGPLVSNLEITHTSNLTCTV  
 KFSNTTYTTNSQCIRWVTPPTQIVCLPSGIFFVCGTSAYRCLNGSSSESMCFLSFLVPPMT  
IYTEQDLYSYVISKPRNKRVPILPFVIGAGVLGALGTGIGGITTSTQFYKLSQELNGDM  
ERVADSLVTLQDQLNSLAAVVLQNRRLDLLTAERGGTCLFLGEECCYYVNQSGIVTEKVKEIRDRIORRAEELR  
NTGPWGLLSQWMPWILPFLGPLAAIILLLLFGPCIFNLLVNEVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPAS  
RSDVNDIKGTPPEEISAAQPLLRPNASAGSS

**FIGURE 4**

- 1) NSLAAVVLQNRRLDLLTAESGGTFLFLEEK
- 2) NSLAAVVLQNRRLDLLTAERGGTCLFLGEEC
- 3) DSLAAVTLQNHQGLDLLTAEGGLCYFLGEDC
- 4) DSLAAVTLQNHQGLDLLIAEKGLCTFLGEEC
- 5) DSLAAVTLQNCRGDLDLLTAEGGHHYTFLEEEC
- 6) LQNRRLDLLFLKEGGLC
- 7) DSLAKVVLQNRRLDLLTAEGGICLALQEK

**FIGURE 5**

TSFVEKANGVKCHKYKLSFHXETTHNYVKSVIYALQEAFRVYLPILPASPTSPSTNKDPPSTQMVQKEIDKRVNSEPKS  
 ANIPQLXPLQAVGGREFGPVHVFPFSLPDLKQIKTDLGKFSNPDGYIDVLQGLGQFFDLTWRDIMSLLNQTLPNER  
 SATITAAXEFGDLWYLSQVNDRTTEEREXFPTGQAVPSLDPHWDTESEHGDWCCRHLTCVLEGLRKRKKSMMNYSM  
 MSTITQGREENPTAFLERLREALRKRAASLSPDSSEGQLILKRKFITQSAADIRKKLQKSAVGPEQNLETLLNLATSVFY  
 NRDQEEQAEQDKRDXXXGHRFSDHPQASGLWRLWKREKLGLNAXXGLLPVRSRTLXKRLSKXXAAPSSMPLISRES  
 LEGPLPQGTKVLXVRSHXPD/SSSRT

**FIGURE 6**



[illegible]

### FIGURE 7

7/64

```

01/      TAAATCCGCATGGCCCTCCCTTATCATATTTTCT
02/      TAAATCCCC-TGGCCCTCCCTTATCATATTTTCT
03/      TAAATCCGCATGGCCCTCCCTTATCATATTTTCT
04/      TAGATCCTCATGGCCCTCC-TTGTATATTTTCT

01/CTTTACTGTTCTTTTA-CCCTCTTTCACTCTCACTGCACCCCTCCATGCCGCTGTATGACC
02/CTTTACTGTTCTTTACCCCTTTCACTCTCACTGCACCCCTCCATGCCACTGCACCCCT
03/CTTTACTGTTCTTTA-CCCTCTTTCTCTCACTGCACCCCTCCATGCTGCTGTACAACC
04/CTTTACTGTTCTTTA-CCCTCTTTCACTCTCACTGAACCCCTCCATGCCACTGTACTACC

01/AGT-----AGCTCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGT
02/GTCCATGCCGCTCTCATGCCAGTAGTCCCTTAGCAAGAGTTTCTATGGAGAATGCAGCGT
03/AGC-----AGCTCCCTTACCAAGAGTTTCTATGAAGAATGCGGCTT
04/AGT-----AGCTCCCATACCAAGAGCTTCTATGGACAATGCGGCTT

01/CCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCTTCACTGC
02/CCCGGAAATATTGATGCCCCATGTATAGGAGTTTATCTAAGGGAACCCCTTCACTGC
03/CCCGGAAATATTGATGCCCCATCAATAGGAGTTTACCTAAGGGAACCTCCACCTTCACTGC
04/CCTGGAATATTGATGACCCATCGTATAGGAGTTTCTAAGGGAACCCCTTTTACCAC

01/CCACACCCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTGCATGCATGCAAACTC
02/CCACACCCATATGCCCCGCAACTGCTATACTCTGCCACTCTTGCATGCATGCAAACTC
03/CCACACCCATATGCCCCGCAACTGCTATACTCTGCCACTCTTGCATGCATGCAAACTC
04/CCACACCTATATGACCC-----

01/ATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGTCACTGTCTGT
02/ATTATTGGACAGGAAAAACGATTAATCCAGTTGTCCTGGAGGACTTGGAG-----
03/ATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCCTGGAGGACTTGGAGCCACTGTCTGT
04/-----

01/TGGACTTACTTCACCCAACTGGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAAGAGA
02/--GACTCACTTCACTCATACAGTATGTCTGATGGGGGTGGAGTTCAAGATCAGGCAACAGA
03/CGGACTTACTTCACCCATACTGGTATGTCTGAGGGGGTGGAGTTCAAGATCAGGCAAGAGA
04/-----

01/AAAACATGTAAAAGAAGTAATCTCCCAACTACCCGGGTACATGGCACCTCTAGCCCCTACA
02/AAAACACATAAAGGAAGTAATCTCCCAACTGACCTGGGTACATAGCACCCCTGCCCCTACA
03/AAAACATGTAAAGGAAGTAACCTCCCAACTGACCCGGGTACATAGCACCCCTAGCCCCTACA
04/-----

01/AAGGACTAGATCTCTCAAACTACATGAAACCTCCGTACCCATACTGCCTGGTAAGCCTA
02/AAGGACTAGATCTCTCAAACTACATGAAACCTCCATACCCATACTGGCTGGTAAGCCTA
03/AAGGACTAGATCTCTTAAACTACATGAAACCTCCATACCCATACTTGCCTGGTAAGCCTA
04/-----

01/TTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGCCCAAACCTACTAACTGTTGGAT
02/TTTAATACCACCCTGACTGGGCTCCATGAGGTCTCGGCCCAAACCTACTAACTGTTGGAT
03/TTTAATACCACCCTCACTGGGCTCCATGAGGTCTCGGTCCAAACCTACTAACTGTTGGTT
04/-----

01/ATGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAACT
02/GTGCCTCCCCCTGCACTTTAGGCCATACATTTCAATCCCTATACCTGAACAATGGAACAACT
03/GTGCCTCCCCCTGTATTTAGGCCATGCATTTCAATCCCTGTACCTGAACAATGGAACAACT
04/-----TGCCTTCAGGCCATACATTTCAATCCCTGTA-----

```

**FIGURE 8.1**

## 8/64

01/TCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATA  
 02/TCAGCACAGAAATAAACACCACTTCTGTTTTAGTAGGTCCTC---TTTCCAATCTGGAAATA  
 03/ACAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTTCCAATCTGGAAATA  
 -----  
 01/ACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTACATACACAACCAACTCCCA  
 02/ACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTATAGACACAGCAACTCCCA  
 03/ACCCATACCTCAAACCTCACCTGTGTAAAATTTAGCAATACTGTAGACACAACCAACTCCCA  
 04/-----  
 01/ATGCATCAGGTGGGTAACCTCTCCACACAAATAGTCTGCCTACCTCAGGAATATTTTTTG  
 02/ATGCATCAGGTGGGTAACCTCTCCACACGAATAGTCTGCCTACCTCAGGAATATTTTTTG  
 03/ATGCATCAGGTGGGTAACCTCTCCACACGAATAGTCTGCCTACCTCAGGAATATTTTTTG  
 04/-----  
 01/TCTGTGGTACCTCAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCA  
 02/TCTGTGGTACCTCAGCCTATCATTGTTTGAATGGCTCTTCAGAATCTGTGTGCTTCCTCTCA  
 03/TCTGTGGTACCTTAGCCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCA  
 04/-----  
 01/TTCTTAGTGCCCCCTATGACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAA  
 02/TTCTTAGTGCCCCCTATGCCATCTACACTGAACAAGATTTATACAATCATGTCATACCTAA  
 03/TTCTTAGTGCCCCC-ATGACCATTTACACTGAACAAGATTTATACAATTATGTTGTACCTAA  
 04/-----  
 01/GCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATAGGAGCAGGAGTGCTAGGTGCAC  
 02/GCCCCGCAACAAAAGAGTACCCATTCTTCCTTTTGTTATTGGAGCAGGAGTGCTAGGCGGAG  
 03/GCCCCACAACAAAAGAGTACTCATTCTTCCTTTTGTTATCGGAGCAGGAGTGCTAGGTGCAC  
 04/-----  
 01/TAGGTACTGGCATTGGCGGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAGAA  
 02/TAGGTACTGGCATTGGCGGTATCACAACTCTACTCAGTTCTACTACAACTGTCTCAAGAA  
 03/TAGGTTCTGGCATTGGCGGTACCACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAA  
 04/-----  
 01/CTAAATGGGGACATGGAACGGGTGCGCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTC  
 02/CTTAAAGGTGACATGGAATGGGTGCGTGATACCTGGTCACCTTGCAAGATCAACTTAACTC  
 03/CTCAATGGTGACATGGAATGGGTGCGCGACTCCCTGGTCACCTTGCAAGATCAACTTAACTC  
 04/-----  
 01/CCTAGCAGCAGTAGTCCTTCAAATCGAAGAGCTTTAGACTTGCTAACCCTGAAAGAGGGG  
 02/CCTAGCAGCAGTAGTCCTTCAAATCGAAGAGCTTTAGACTTGCTAACCCTGAAAGAGGGG  
 03/CCTAGCATCAGTAGTCCTTCAAATCGAAGAGCTTTAGACTTGCTAACCCTGAAAGAGGGG  
 04/-----  
 01/GAACCTGTTTATTTTTAGGGGAAGAATGCTGTTATTATGTT-----  
 02/GAACCTTTTTATTTTTAGAGGAAAAATGCTGTTGTTATGTT-----  
 03/GAAGCTGTTTATTTTTAGGGGAAGAATGTTGTTATTATGTTATTTTAGCGGAAGAATGTTGT  
 04/-----  
 01/-----AATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACA  
 02/-----AATCAATCCGGAATCATCACCGAGAAAGTTAAAGAAATTCAGGTCGAATATA  
 03/TATTATGTTAATCAATCCTGAATTGTCACAGAGAAAGTTGAAGAAATTCGAGATTGAATACA  
 04/-----  
 01/ACGTAGAGCAGAGGAGCTTCGAAA-CACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCT  
 02/ACGTAGAGCAAAGGAGCTGCAAAA-CACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCT  
 03/ACGTAGAACAGAGGAGCTTCAAAAACACCAGACCCTGGGGCCTCCTCAGCCAATGGATGCCCT  
 04/-----

FIGURE 8.2

9/64

01/GGATTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCTGTGTA  
02/GGATTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATTGTTACTCCTCTTTGGACCTGTGTA  
03/GGATTCTCCCTTCTTAGGATCTCTAGCAGCTCTAATATTGATACTCCTCTTTGGACCTGTGTA  
04/-----

01/TCTTTAACCTCCTTGTTAACTTTGTCTCTTCCAGAATCGAAGCTGTAAACTA-----  
02/TCTTTAACCTCCTTGTTAAGTTTGTCTTTTCCAGAATCGAAGCAGTAAACTACAAATCGTTC  
03/TCTTTAACCTCCTTGTTAAGTTTGTCTCTTCCAGAATCAAAGTTGTAAAGCTACAAATCGTTC  
04/TCTTTAACCTCCTTGTTAAGCTTGTCTCTTGCAGAATCGAAGCTGTAAACTACAAATGCTTG

01/--CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCCCTGGACCGGCCTG  
02/TTCAAATGGAGCCCCAGATGCAGTCCATGAGTAAATCTACCACGGACCCCTGGACCGGCCTG  
03/TTCAAATGGAACCCAGATGAAGTCCATGACTAAGATCTACCGTGGACCCCTGGACCGGCCTA  
04/TTAAATAGAGCCCCAGATGCAGTCCATGGCTAAGATCTACCACGGACCCCTGGACCGGCCTG

01/CTAGCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCCTGAGGAAATCTCAGCTGCAC  
02/CTAGCCCATGCTCTGATGTTAATGACATCAAAGGCACCCCTCCCGAGGAAATCTCAACTGCAC  
03/CTAGCCCATGCTCCAATTGTAATGATATCGAACGCACCCCTCCCGAGGAAATCTCAACTGCAC  
04/CTAGCCCATGCTCTGATGTTGATGACATTGAAGGCACGGCTTCCGAGGAAATCTCAACTGCAC

01/AACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCC  
02/AACCTCTACTACGCCCCAATTCAGCAGGAAGCAGTTAGAGTGGTTGTTGGCCAACCTCCCC  
03/AACCCCTACTATGCCCCAATTCGCAGGAAGCAGTTAGACTGGTCGTCAGCCAACCTCCCC

04/GACCCCTACTACACCCCAATTTAGCGGGAAGCAATTAGAGCAGCCTATGGCCACCTCCCC

**FIGURE 8.3**

10/64

CTTCCCCAACTAATAAGGACCCCCCTTCAACCCAAACAGTCCAAAAGGACATAGACAAAGGA	3
CTTCCCCAACTAATAAGGACCCCCCTTCAACCCAAACAGTCCAAAAGGACATAGACAAAGGA	4
CTTCCCCAACTAATAAGGACCCCC-TTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGG	5
CTTCTCCAATAATAAGGACCCCC-TTCAACCCAAATGGTCCAAAAGGAGATAGACAAAAGG	6
CTTCCCCAAATAATAAGAACCCCC-TTCAACCCAAACGGTCCAAAAGGAGATAGACAAAGG	7
GTAAACAATGAACCAAAGAGTGCCAATATTCCCTGGTTATGCACCCCTCAAGCGGTGGGAG--	3
GTAAACAATGAACCAAAGAGTGCCAATATTCCCTGGTTATGCACCCCTCAAGCGGTGGGAG--	4
GTAAACAGTGAACCAAAGAGTGCCAATATTCCCAATTATGACCCCTCAAGCAGTGGGAGGA	5
GTAAACAATGAACCAAAGAGTGCCAATATTACACGATTATACTCGCTCCAAGCAGTGGGAG--	6
GTAAACAATAACCAAAGAATGCCAATATTCCCGATTATGCCCCCTCAAGCGGTGGGAG--	7
A-AGAATTCGGCCCGAGCCAGAGTGCATGTACCTTTTTCTCTCTCAC-ACCTGAAGCAAATTA	3
A-AGAATTCGGCCCGAGCCAGAGTGCATGTACCTTTTTCTCTCTCAC-ACCTGAAGCAAATTA	4
AGAGAATTCGGCCCGAGCCAGAGTGCATGTGCTTTTTCTCTCCAG-ACCTAAAGCAAATAAA	5
-GAGAATTTGGCCCGAGCCAGCGTGCATGTACCTTTTTCTCTCTCAG-ATTTAAAGCAAATTA	6
-GAGAATTCGGCCCGAGCCAGAGTGCACGTACCTTTTTCTCTCTTAGACTTTAA--TTAA	7
ATAGACNTAGGTNAATTNTCAGATAGCCCTGATGGYTATATTGATGTTTTACAAGGATTAGGA	3
ATAGACXTAGGTXAATTXTCAGATAGCCCTGATGGXTATATTGATGTTTTACAAGGATTAGGA	4
ACAGACTTAGGTAAATTCTCAGATAACCCCTGATGGCTATATTGATGTTTTACAAGGTTAGGA	5
ATAGACCTAGGTAAATTCTCAGATAACCCCTGATGGCTATATTGATGTTTTACAAGGTTAGGA	6
ATAGACCTAGGTAAATTCTCAGATAACCCCTAATGGCTATATTGATGTTTTACAAGGTTAGGA	7
TTCTGAGTTCCTGCACTAACCTCAAAT	1
CAATCCTTTGATCTGACATGGAGAGATATAATATTACTGCTAAATCAGACGCTAACCTCAAAT	3
CAATCCTTTGATCTGACATGGAGAGATATAATATTACTGCTAAATCAGACGCTAACCTCAAAT	4
CAATCCTTTGATCTGACATGGAGAGATATAATGTCACTGCTAAATCAGACCTAACCCCAAAT	5
CAATCCTTTGATCTGACATGGAGAGATATAATGTTACTGCTAAATCAGACCTAACCCCAAAT	6
CAATCCTTTGATCTGATATGGAGAGATATAATGTTACTGCTAAATCAGACCTAACCCCAAAT	7
GAGAGAAGTGCCGCCATAACTGCAACCCAAGAGTTTGGCGATCCCTGGTATCTCAGTCAGGTC	1
GAGAGAAGTGCTGCCATAACTGGAGCCCGAGAGTTTGGCAATCTCTGGTATCTCAGTCAGGTC	3
GAGAGAAGTGCTGCCATAACTGGAGCCCGAGAGTTTGGCAATCTCTGGTATCTCAGTCAGGTC	4
GAGAGAAGTGCCACCATAACTGCAGCCTGAGAGTTTGGCGATCTCTGGTATCTCAGTCAGGTC	5
GAAAAAAGTGCTGCCATAACAGCAGCCTGAGAGTTTGGCGAACTCTGGTATCTCAGTCAGGTC	6
GACAGAAGTGTCGCCGTAACCTGGAGCCCGAGAGTTTGGCAATCTCTGGTATCTCAGTCAGGTC	7
AATGACAGGATGACAACAGAGGAAAGATAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGT	1
AATGATAGGATGACAACGGAGGAAAGAGAACGATTCCCCACAGGGCAGCAGGCAGTTCCCAGT	3
AATGATAGGATGACAACGGAGGAAAGAGAACGATTCCCCACAGGGCAGCAGGCAGTTCCCAGT	4
AATGATAGGATGACAACAGAGGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGT	5
AATGATAGGATGACAACAGATGAAAGAGAATGATTCCCCACAGGCCAGCAGGCAGTTCCCAGT	6
AATGATAGGATGACAACAGAGGAAAGAGAACGATTCCCCACAGGCCAGCAGGCAGTTCCCAGT	7
GTAGACCCCTCATTAGGACACAGAATCAGAACATGGAGATTGGTGCCGAGACATTTGCTAACT	1
AACT	2
GTAGCTCCTCATTGGGACACAGAATCAGAACATGGAGATTGGTGCCGAGACATTTACTAACT	3
GTAGCTCCTCATTGGGACACAGAATCAGAACATGGAGATTGGTGCCGAGACATTT	4
CTAGACCCCTCATTGGGACACAGAATCAGAACATGGAGATTGGTGCTGCAGACATTTGCTAACT	5
GTAGACCCCTCATTAGGACACAGAATCAGAACATGGAGATTGGTGCCACAGACATTTGCTAACT	6
GTAGACCCCTCACTGGGACACAGAATCAGAACATGGAGATTGGTGCCGAGACATTTGCTAACT	7

FIGURE 9.1

## 11/64

TGCGTGCTAGAAAGGACTAAGGAAAAGCTAGGAAGA----	1
TGCGTGCTAGAAAGGACTAAGGAAAAGCTAGGAAGA---CTATGAATTATTCAATGATGTCCACT	2
TGCGTGCTAGAAAGGACTAAGGAAAAGCTAGGAAGA---CTATGAATTATTCAATGATGTCCACT	3
TGTGTGCTAGAAAGGACTAAGGAAAAGCTAGGAAGAAGTCTATGAATTACTCAATGATGTCCACA	5
TGCGTGCTAGAAAGGACTAAGGAAAAGCTAGGAAGAAGCCCATGAATTATTCAATGATGTCCCT	6
TGCGTGCTAGAAAGGACTAAGGAAAAGCTAGGAAGAAGCCTGTGAGTTATTCAATGATGTCCACT	7
ATAACACAGGGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG	1
ATAACACAGGGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG	2
ATAACACAGGGGAAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG	3
ATAACACAGGGG-AAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG	5
ATAACACAGGGG-AAAGGAAGAAAATCCTACTGCCTTTCTGGAGAGACTAAGGGAGGCATTGAG	6
ATAACACAGGGG-AAAGGAAGAAAATCCTACCGCCTTTCTGGAGTGACTAACGGAGGCATTGAG	7
GAAGCATACC---AGGCAAGTGGACATTGGAGGCTCTGGAAGGGGAAAAGTTGGGAAAAGTA	1
GAAGCATACC---AGGCAAGTGGACATTGGAGGCTCTGGAAGGGGAAAAGTTGGGCAAATTG	2
GAAGCATACC---AGGCAAGTGGACATTGGAGGCTCTGGAAGGGGAAAAGTTGGGCAAATTG	3
GAAGCGTGCC232AGGCAAGTGGACTTTGGAGGCTCTGGAAGGGGAAAAGCTGGGCAAATTG	5
GAAGCATACC238AGGCAAATGGACTTTGGAGGCTCCAGAAAAGGGGAAAAGCTGAGCAAATTG	6
GAAGCATACC233AGGCAAGCGGACTTTGGAGGCTCTGGAAGGGGAAAAGCTAGGCAAATCA	7
TATGTCTAATAGGGCTTGCTTCCAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCC-AA	1
AATGCCTAATAGGGCTTGCTTCCAGTGCAGTCTACAAGGACGCTTTAGAAAAGATTGTCC-AA	2
AATGCCTAA	3
AATGCCTAATAGGGCTTGCTTCCAGTGCAGTCTACAAGGACACTTTAAAAAAGATTGTCC-AA	5
AATGCCTAACAGGGCTTGCTTCTAGTGTGGTCTACAAGGACACTTTAAAAAAGATTGTCC-AA	6
AATGCCTAATAGGGTTTGCTTCCAGTGCAGTCTACAAGGACACTTTAAAAAAGATTGTCCAA	7
-TAGAAATAAGCCACCACCTCGTCCATGCCCCCTTATGTCAAGGGAATCACTGGAAGGCCCACT	1
GTAGAAATAAGCCGCCCTCGTCCATGCCCCCTTATGTCAAGGGAATCACTGGAAGGCCCTACT	2
GTAGAAATAAGCCGCCCTCGTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACT	5
GTAGAAACAAGCTGCCCCCTTGTCATGCCCCCTTATGTCAAGGGAATCACTGGAAGGCCCACT	6
-TAGAAATAAGCCGCCCTCGTCCATGCACCTCGTGTCAAGGGAATCACTGTAAGGCCCACT	7
GCCCCAGGGGATGAAGGTCTCTGAGTCAGAAGCCACTAACCAGATGA	1
GCCCCAGGGGACGAAGGTCTCTGAGTCAGAAGCCACTAACCTGATGA	2
GCCCCAGGGGACAAAGGTCTCTGAGTCAGAAGCCACTAACCAGATGA	5
GCCCCAGGAGATGAAGGTCTCTGAGTCAGAAGCCACTAACCAGATAA	6
GCCCCAGGGGACGTAGGTCTCTGAGTCAGAAGCCACTAACCAGATGA	7

**FIGURE 9.2**

12/64

RTPLSTQTVQKDIDKGVNNEPKSANIPWLCTLQAVGEEFGPARVHVPPFSLSHLKQIKIDG SDSPOG  
- = === ----- == ===== = ===== ===== ===== = = == ===  
KDPPSTQMVQKEIDKRVNSEPKSANIPQLPLQAVGGREFGPARVHVPPFSLPDLKQIKTDLGKFSDNPDG  
  
YIDVLQGLGQSFDLTWRDIILLNQTLSNERSAAITGAREFGNLWYLSQVNDRTTEERERFPTGQQ  
===== ----- ===== ----- ===== ===== =====  
YIDVLQGLGQFFDLTWRDIMSLLNQTLPNERSATITAAXEFGDLWYLSQVNDRTTEEREXFPTGQQ  
  
AVPSVAPHWDTESEHGDWCRRHLLTCVLEGLRKTRK TMNYSMMSTITQGK  
----- ===== -----  
AVPSLDPHWDTESEHGDWCCRHLLTCVLEGLRKTRKSMNYSMMSTITQGR

**FIGURE 10**

13/64

[illegible]

**FIGURE 11**



14/64

[illegible]

**FIGURE 12**

15/64

agttgcaatttccttgccctcaactctgagagaaaacccagccacatctccagcaaaacaaga  
|||||  
agttgcaattccttgccctcactgtgagacaaacccagacacatctccagcacacaaga 2299

acttcaaaacacctgaactgcagcagccagggttctctccaggaccctccccaggat  
|||||  
acttcgaaatgcctcaacctcaggtgccagggttctctccagaaccttctccccaggag 2359

cttgcttcaagtgccggaaatctgaccattggggccaaggaatgcctgcagcccaggattc  
|||||  
cttgctacaagtgccgaaatctggccactggggccaaggaatgccacagaccaggattc 2419

ctcctaagccacgtcccatttgtgcaggacccactggaaatcggactgtccaactcacc  
|||||  
ctcctaagctgtatcccatctctgtgggacccactaaaaatcagactgttcaactcacc 2479

cggcagccaatcccagagcccttggaactctggccaaggctctctgactgactccttcc  
|||||  
tggcagccacttccagagcccttggaactctagccaaggctctctgactgacccttct 2539

cagatcttctcggcttagcagctgaagactgacactgcccgatcacttcagaagtccct  
|||||  
gagatcttcttggcttagcagctgaagactgacactgccagatcgctcggaagcctaca 2599

ggaccatcacggatactgagcttcaggttaactctcacagtggaggctaagtcctatccct  
|||||  
ggaccatcacagat----gctccaggttaactctcacagtagagggttaagctctgtccct 2654

gtttaatcgatacaggggctaccactccacatcaccttcttttcaagggcctgtttccc  
|||||  
tcttaatcaatatggaggctaccactgcacattaccttcttttcaagggcctgtttcct 2714

tttccccataactgttgtgggtattgacggccaagcttcaaaaccccttaaaactcccc  
||  
ttgctccataactgttgtgggtattgacggccaggcttcttaacctcttaaaactcccc 2774

cactctgggtgccaacttggaacaactctttttatgcactctttttcagttatcctcacct  
|||||  
aactctagtaccaacttagacaatactcttttaagcactcctttttagttatccccactt 2834

gccaggttcccttattaggccgagacattttaaccaaattatctgcttccccgactattc  
|||||  
gccaggttcccttattgaggccgagacacttcaactaaattatctgcttccctgactattc 2894

ctgggttacagccacatctccttgccgcccttcttcccaacccaaagcctccttcatatc  
|||||  
ctggactacagctacatctcatgtctgcccttcttcccaatccaaagcctcctttgcatc 2954

ttcctctcatatccccccaccttaaccacacaagtatgggacacctctactccctccctgg  
|||||  
ttcctgt--atcccccaaccttaaccacacaagataagataacctctattccctccttgg 3011

**FIGURE 13.1**

[illegible]

**FIGURE 13.2**

17/64

atataaactcacaaaaggaaacctagctgaccccatagattctaaatcctttccccactc  
 |||||  
 atataaactcacaaaaggaaacctagttgaccccatagatcctaaatcgttccccactc 3898

ctctttccattccttgaagacagctttagagactgctccacactagctctccctgtctc  
 |||||  
 ctctttccattccttgaagacagctttagagactgtctccactctagctctccctgactc 3958

atcccaacccttttcattacacacagccgaagtgcagggctgtgcagtcggaattcttac  
 |||||  
 atcccaacacttttcattacacacagctgaagtgcagggctgtgcagtcagaattcttac 4018

acaaggaccgggacatgacctgtagcctttttgtccaaacaacttgaccttactgtttt  
 |||||  
 acaaggaccgggatcgcatcctgtagcctttttgtccaaacaacttgaccttactgtttt 4078

aggctcgccatcatgtctccatgcggtagcttccgctgcccataacttttagaggccct  
 |||||  
 aggctggccatcatgtctccatgcagcgtctgtgccaccctaataacttttagaggccct 4138

caaaatcacaaactatgctcaactcactctctacagctctcacaacttccaaaatctatt  
 |||||  
 caaaatcacaaactatgctcaactcattctctacagctctcataatttccaaaatctatt 4198

ttctttctcacacctgacgcataactttctgtccccggctccttcagctgtattcact  
 |||||  
 ttcttctcacacctgacacataactttctgtccccggctccttcagatatactcact 4258

ctttgttgagtctcccacaattaccattcttctgcccagacttcaatctggcctccca  
 |  
 c--catttattctcccacaattaccattattctgcccagacttcaatccggcctccca 4316

cattattctggataccacacctgacctgatgattgtatgtctctgatctacctgacatt  
 |||||  
 cattattctggataccatactgacctcatgactgcattctctgatccacctgacgtt 4376

caccatcttccccatatttcttcttctgttctcatgttgatcacatttggtttac  
 |||||  
 caccatcttccccacatttcttctgcccgtgttctcaccctgatcacacttggtttat 4436

tgacggcagttccaccaggcctgatcgccactcaccagcaaaggcaggctatgctat  
 |||||  
 tgatggcagttccaccaggcctaatacgccactcaccagcaaaggcaggatagctat 4493

gaactgattgccttaactcgggccttcaactcttgcaaagggactacacgtcaatatttat  
 |||||  
 gaactagttgccttaattcaagccctcactcttgcaaagggactacgtgtcaatatctat 4553

actgactctaaatatgccttccatatcttgaccaccatgctgttatatgggctgaaaga  
 |||||  
 actgattctaaatatgccttccatatcttgaccaccatgcggtcatatgggctgaaaga 4613

ggtttctcactacgaagggtcctccatcattaatgcctctttaataaaaaactcttctc  
 |||||  
 ggtttctcactacacaagtgtcctccatcattaatgcctctttaagaaa-ctctgctc 4672

FIGURE 13.3

aaggctgctttacttccaaaggaagctggagtcacacactgcaagggccaccaaaggcg  
|||||  
aaggctgctttacttccaaaggaagctggggtcattcactgcaaggggcatcaaaagact 4732

tcagatccattactctaggaatgcttatgctgataaggtagctaaagaagcacctagc  
|||||  
tcagatccattgctctaggcaatgcttatgctgataaggtggctagacaagcagctagc 4792

gttccaacttctgtccctcatggccagtttttctccttcccatcagtcattcccacctac  
|||||  
tctccaacttttgtccctcatggccagtttttctccttcacatccgtcactcccacctac 4852

tccccattgaaacttcgcctatcaatctcttctcacacaaggc aaatggttcttagac  
||| |  
tccacagctgaaacttcacctatcaagctcttccccgcaaggtaaatggttcttagac 4912

caaggaaaaatatctccttcagcctcacaggccattctattctgtcatctttcataac  
|||  
caaggaaaaatatctccttcagcctcacaggccattctattctgtcgtcatttcataac 4972

ctcttccatgtaggttacaagccactagtccacctcttagaacctctcatttccctt-cca  
||  
cttttccatgtaggttacaagccactagcctgtctcttaggacctctcatttcccttcca 5032

tcgtggaacatatcctcaaggaaatcacttctcagtgttccatctgctattctactacc  
||  
tcatggaaatctatcctcaaggagatcacttctcagtgttccatctgctattctgctacc 5092

cctcagggattgttcaggccccctccctccttacacatcaagctcggggatttgccct  
|||  
cctcagggattgttcaggcctcctcccttctacacataaagctcggggatttgccct 5152

gccaggactggcaaatgactttactcacatgcctgagtcaggaaactaaaaacctc  
|||  
gccaggactggcaaatgactttactcacatgcctcgggtcagaaaactaaaaatatctc 5212

ttggtctgggtagacactgtcactggatgggtagaggcctttcccacagggctctgagaag  
||  
ttagtctgggtagacactttcactgggtgggtagaggcctttcccatagagtctgagaag 5272

gcactgcagtcatttcttcccttctgtcagacataattccttggttggccttcccacc  
|||  
gccaccgcggctatttcttcccttctgtcagacataattccttggttggccttcccttc 5332

tctatacagtc caataacggagcagcctttattagtcaaatcacctgagcagttttcag  
|||||  
tctatacagctctgataacggaccagcctttactagttaaatcacccaagcagtttctcag 5392

gctcttggtatttcagtggaaaccttcgtacccttactgtcctcaatcttcaggaaaggta  
|||  
gctcttggtatttcagtggaaaccttcatacccttaacatcctcaatcttcaggaaaggta 5452

gaatggactaatggtcttttaaaaaacacacccaccaaactcagcctccaacttaaaaaag  
||  
aaaccgactaatggtcttttaaaagacacacctcaccaagctcagcctccaacttaaaaaag 5512

**FIGURE 13.4**

19/64

TGCCTTTATTTCCGTAGGCTGGTCATATGGCGCTAGCACTCACATAAAGCTACCGAGGAG  
 AGCGAATGAAACCAAAATCACTTTACCTTCACAGCACGAGGCCGTGCTCCCTCTCGATAT  
 TTGGCCCGTGTGTGCGATACCGCCCTCTGGACGTGGTGATCAAATAAACTCCCTAGCTCC  
 CCGCCGCTCGACGCCATCTTGCTACTTTGATCCTCGCAGGGAGGACAACATCCGCCCTA  
 CTGAGCTCCCTTTTATCCAATAAGAGAGCGGGATGAGTTAAGGAGTGCCAGGATTGGCTG  
 GAGAATCGACAGCGTCGCCCATCGTTTCCTGCGTGCGAAGATTGATGAACGAGGTGCCG  
 CCCCCGAGCGGCTCGGCGGAGAGGCGCGGTGGGTGACAGAAGCTTTCTTGTCACCCAC  
 TACAGGCTTACGGCAGGATGCGCAGCGGGAGAGGGGCGGGCCGAGGGGGCGGGGCC  
 GATCGATCTCCTCCGGCTCCGACGTCTCGGCCTGCGGGTCCCGGGTCTTTGCGGCGC  
TAGGGTGGGCGAACCAGAGCGACGCTCCGGGACGATGTGGGGCAGCGATCGCCTGGCGG  
GTGCTGGGGAGGCGGGCGGCAGTGACTGTGGCCTTACCACGCTCGCGACTGCTTCC  
TCCACCTGCGCGGCGTCTCGTGGCCAGCTGCATCTGCTGCAGGTAACCTGCCGCCCC  
 GAGCCACCTGATCTTACGCTGGGGTCCGACGAGGCCGAAGCTCTCAGGGACGCGGCGG  
 GACACCGGCTGCCACCCGGGCGCGCCGAAGCGCGCAGAGATCAGGGTCCCTCGACGGCA  
 GGGCCCTTCTGGGTAGTCTCTGGATCCACAAGTCCAGTGACGCCCTGGGCTCGTCTTAT  
 CCCAGGTCTTTTCACTTTGGTGAACTGAACCTAGAAACGTCCTAATATTCTACCACTGTT  
 TTTATAAATATTCTTATTCCAGGCTGGAAAAGCTCCTGAGAAGTGGTTTGTATTTATTA  
 TTTTAAAAGGTGTTTTCTTGCCAGCCATTTCAGTTAACCTGCGCTGCTGCCGTCCGGG  
 CCGCGAGAGCGGGACGCGAGAGTTGTTGGCGGAGCCCTGTGCGTTCCCGGGGACTAAGCA  
 CCGCGTCCCATGAGCGGGAAAGGTTAATACAATGATGGTTCTGCCCTGCGTGCCTGACGC  
 GGAACACAGCTGTAGTGTGTTAGGAACACATAACGTAGTTAAGATCACTTGAAGCTCTGC  
 GATCAGTCGCCCTTCTGGACGTTGTGGTTAGGATGTTTACAGTTCTAACCAGTGGTGGA  
 GATACAGCGCTCCATATTTTATAATTAAAAATAGAGGCACATGGTCTCACGAGTTTGAGT  
 GTACTTATGGGGGCAAAAGGACGGCGTATTTGAAATCCTCATAAATCCTGGATGCATGGT  
 ACCCACCAGTGGCTAATCTATGCAATGAATAGAGTTTGAATAATTTCAAGCATCCCTTC  
 TTTCCACTTGAGTTACTTCCCCATACCTAGGGGAAGATATTTTGGTCCACTGAAAACAT  
 GAGTTTCAGCAGAATCCTCCTATCATCGTCGTTATTATTTTTTACCCTAAGTAGACAATC  
 TTTTGGTTTTTGATGGGCTTTATGGCTAGAGACAAATCAGTCACTGTCAACCAAGTTCAG  
 GTAGAAGTTGGTTCACTGCTCTGTGAGCTTCGATGGGATTTTCAACATGTTTTCAAATC  
 TGCACTTAATAGTAGGAATGCTTTCTTACAGTAACCTCTAATTTGATCCTAAGATGTAGTT  
 GTTACCTTACATTTCATCACTGTTTAAAGAAATTTAGTGGTCTTGATCTTTGTTTTAAATTTT  
 GAGCCTTCGGGAAGTACTTATAAGAATTAATTCATGCATATCTTTTGAATGTAAATGT  
 CTTTAGCCCTGGAACAAATTTGCTGTTTCTGTTTACGCCATATTAGCAGAATAGGTCAACT  
 TTACTTTCTAATTATCAATGTAATAAGTTTATTACTTTATAGATTCCATAAATCTATACA  
 TTTATTCTCGATGAATTATATAAATTTATAGAATTTATGTTTTATAGAAAATTTGAAA  
 GCATGGAATTTATTAACAAGAAAATAAGTTACCCATAATCCAGAACTTAGAGGTGACT  
 AATGTTGACAGTTTGGATCAAATCTTCCAGTTTGTCTTCTAATCTTTATTTTTAACATAA  
 ATGAGGTCTGTATACACACGTACAGTTTGTGTCCTGGTGTTTTTATTTAATGTTATTA  
 TGAGTGTTTTTATTTTGTAAAAGGTCATCATTTTAAAGTTGTTAATTAGTATTCTAGCACA  
 AATTTGCCATAATTTATTTAATTGTTTACTATGATTGACCATTTAGATTGTACTTAATTT  
 TTAGGCATTAGAAGTGATAAATATATTTAATCAGACGTTGAAAAAATACACATCTTTGT  
 TTAGAAAACATCATTTTATTTCTGGTTGTCTAGGATAGATTCCAGAAATCTTGGGTTAG  
 AGGCCATAGATAATTATGAAAGCAGAAAGATTCACAAGTTGGGAGTTAATACTTGAATTA  
 CTTTATTTGGGGTGAGCATTGAGTGCATAATACAGATCATGCAGTAATGGGAAGAAGGG  
 TTGGAACAATGGTTTTCTGGCCTATGTACAGCTTACCTTGAAGCTTTTAAAGATACAGAT  
 GTTCTGATCAACCTCAGACCTATTAAATCAGACCTAAAATCTTAGGGAATAGGCTTTAG  
 GCATCTCTAATTTTAAAAAATTTATTCAGGCTACTTGGATGCACAAAAGAGTTGAGACCT  
 ACTGTCTTAGAATCATAGAATTTAATGACGATAGAGACCTTAAGCATCTAGGTGCTTTC  
 TGTACTTTTACATGTAAGGAACTGGCATTCTAGGCCAGTACCATTGCCATGCAGCTAA  
 TTTGCCCTCTTGTCTATAGCTCACTCTGCATCACCACCTACCGTTCTCACTGTTTCTT  
 CTATAACCAATCTCCTTCCCACTTCTGTCTTCTTACTCATGCCATTCTTCCCTCAGTCAT  
 TTTTCTTCTTCCATACAAATTCATGTCTTTAAAAAGGAATAATCCTACCTCCTCCACA

FIGURE 14.1

20/64

TAGCTTTCCAATTCTCTGTTGCCACATTTGTCTCCCTTTCAATACTTCTCTGTTGTGTT  
ATGTGACACATCACATTTGATATACTCTGTACTGTGTTTCAAGTATTGTATTCTCTGTGTT  
TACTCAAGTCATTATTTTCAGGACTGACTACCCAGTAGATGCTTTAAGTCAGGATTTCTCA  
ACCTTGGCACTGTTGACATTTTGAGCTGGATAATTTTTGTTTTGGGGGCTCTCCTGTAC  
ATTTTAAGATGTTTAACAGCACCTTGGCCTCTATCCAGTAGACGCTGTACTGCCTCCC  
CCTATCTGTGACAACCAAAAGGTCTTCAGACATTGTGAGATGTCTACTGAAGACAAAA  
TCACCTCTGGTTGAGAACCACCGCTTCAACTAAGTTATCTTCTCTGTACTCAGAAGTTGA  
TGTGATTGCAGCAGGGGGAGAGGATTATATACACAGTGAATGCAAACGAACCTAAATCA  
CCATTCCGATATGGCCACACAATTTTCATTTCCCTTGTGTTAGCAAGAGATACCCTAGGC  
TTTGGACCTGATATTTCCCTAAGGCATTCTGATGTATGGTTTTACCTGCAGATTTCCCTGGT  
AATACTGATACCTCAGTTTGGGTCAAAGAAGGTCAATTAATTGATTGATTGATTGACT  
CCTGGAAAAGACGCTCCTTTCTAGCTGTCTCTTTCTCTTTTACCTGAATAGCCAGGGC  
TCTGTGGTTCAAGTGAAGTATTTTGACATAAAAATTAACCTAGAACATTGGTCTGCAGAG  
TTTGTCTCAATATAACTGAGCACATATTGTGGCTTTATGGAGCTGGTTACTACTTTTTGAC  
CAAATAAATAATTAGAAGTATTTTTCTCCTCAATAAGGTTCAATTTTCTTTTTTTCAGT  
GAGCTGGTAGAGTTTCTTTTTTGTATTTTTCAGGGCATCTTTCATATTTCCATCTCTTAA  
GTTTTCTTCAATGAAAGTAGAATTTATCTGGATTATGTATTGCTGACTCTGATGAAAACCC  
ATAGAAAGCATCTGGGGCTTGATCACCTTCATTCTTGAATAGCTCACACGGTTACAGCT  
GATATGGTAACTTAAGACTTTTGATTCCAAATCTAGGCAAAATACACTCAGTTGAAAGAA  
TTTGTGAGCCAGAACAGTTGGACTGTTCTGTGAAAATTTGTGAGAAAAATACACAACATAA  
GTGATACATGATGATGGCTTTCTTAAATATAAAATTGTAATAACATGGTTAATTTCCAGT  
ACGTTATATTGTCCCAGAAGTGGCTCCAACATTGTTTGAAATTTGTCTCATTTAAAGAAA  
CATAAGCTGGCTATGGTGGCTCACGCTGTAAATCCAGCACTTTGGGAGGCTGAGGCAGG  
CAGATCACCTGAGGTGAGGAGTTTCGAGACCAGCCTGGCCAACATGGTAAACCCCATCTC  
TACTAAAAATACAAAAATTAGCCGGGCATTGGTGGGGGCTGTAATCCAGCTACTTGG  
GAGGCTGAGGCAGGAGAAATTGCTTGAATCTGGGAGGTGGAGGTTGCAGTGAGCCGAGATT  
GTGCCACTGCCCTCCAGCCTGGGTGACAGAGTGAGTCTCCGTCTCAAGAAAAA  
AAAAAGCAAGAAACATAAAGACTGGGCATGTTGGCTCATGCCTGTAATCCAGCACTTTGA  
GAGACTGAGGTGGGAAGATCACTTGAGCCCAGGAGGTTAAGGCTGCAGTGAGCCGTGATT  
TTGCCACTGTACTCGAGCCTGGGCAACACAGTGAGATCCTGTCTCAGGAAAAA  
GCATGTAAATGAATGAATTTGATATTTAATATTTTAAATATGAAAACTGTTCTGTAGAG  
ATGTAGATCTTGCCATGTTGCCAGGCTGGCTTTGAACTTCTGGGCTCAAACAATCCTCC  
TGTCTCAGTCTCCCAAGTATAAAGATTACACATGTGAGCCACTGCACCTGGCCTAATAT  
TTTTAACTTAATGAATTTATTTTGATATAAATAAATTAATAACACTGAAGCTTCCTGATA  
TAATAAGTCTTTTTGTGTGTGTGACGGGTTCTCACTCTGTTGCCAGACTGGAGTGTAAT  
GGCACTATCATGGCTCACTGTAGCCTCAACCTCCCTGACTCAAGTGATCCTCCCACCTCG  
GCTTCCTGAGTAGATGGGACCACAGGCGTATGCCACCACACCTGGCTGATTTTAAAT  
TATTATTGATACATATTAATAAAATTATTTTATTTTAAATATGATATATGTGGCTGGGC  
ATGGTGGCTCATGCCTGTAATCCCGACAGTTTGGGAGGCGAGGTGGGAGGATCACTTGA  
GACCAGGAGCTTAAGACCAGCCTAAGCAACATAGTGAGATCCCATCTCTATAGAAAAA  
AAATGGCTAGGTGTGGTGGTGTATGCCCTATATCCAGCTACTCAGGAGACTGAGGTGAG  
AGGATTGCTAGAGCCCAGGAGTTCAAGTTACAGTGACCTATGATTGTGCCAGTGCACTC  
CAGCCTGGGCAACAGAGCAAAATCCTGTCTCAAAAAAAGGTTGCAAAATGCTTAT  
GATGCAATATAAGTAGTGAAAAAGGATATTAATTTGTGCCTATATGAACACAACATATG  
AAAAACTTGACATAGAGAAAAAGGATTAACAAGAAATAGACCAAATTGTTACATGGTTG  
TCTTGTGTTGTGGAGAGAAATACAGTAGTTCATTTGTTTCTTCCAAGTTTATATGTTTC  
CGAGGTCTCTATAATGAGTTTGTAAATTGTTTAAATCATAGAAAACCTTTTTTGGTCCTTG  
GCCACAACTTACATGTTTAAATGTAATTGCTTTTTTAAATGAGAATAAATGTTATATTT  
GCTTTTTTAAACCTATATCCCATAGTTATATGAGCCCTTACAATTATTAAGAGGCTGC  
ATAATATAACGTTTCTGGAAGGGTACAGAAGAAACAGCAGTAATTACCTCTGAGAACAGA  
GACATGGCTTACATTTTACCCTTTTGTACGTTTGTGCTTTTGCCACATGCATTTATTA  
TTCTTCCAATAAATAAGTAAATAAATATGGATTGTATACTCCATCTGGTGGTGTTCAT  
AATCTAAAAATTATATTGCTACATTTTAAAGATGATATGTGTTCTACTTATTAACGTA

FIGURE 14.2

TATGTTAAAATAGTAAATTTATATCTTATTTAATAAATTTCCCTATTGATAGACATTTAAG  
ACAGTCTCAAGTGTTCACTATCATAGAAAATACTGCACAGATAGCTTTTGCTATAGTTTC  
TTTTTTCTTTGAATCGTTAATTGGGAATAAATGCCTCAAATAGTTATATGTGGCTCAACTG  
CTATTTAAGTTTATTGACTGACTGCTGCCATTTTGAATTCGAAGGGGTTGATTAAATTT  
ATAATGCTGCCATAAGAATAAAGGGTATTGGCTTCATTAGCATCCACCAGCATTGGGTG  
TTGGAAATGATTATAGATTATTTAAATGCTCAACAACAAATGTAGATAACAGAGAACTACTA  
TAGAATCTTTTTTTGGACATGTGGAATGTGAATAAGTTTATTTTTCTATGTGAATCCAGAAA  
AATGTATACGAAAAACCTTTTTCTCTCATTTCCTATATGAATAGAATCAAGCTATAGAA  
GTGGTCTGGAGTCACCAGCCTGCATTCTTGAGCTGGGTGGAAGGCAGGCATTTTAGTGAT  
GGGGGACAGGTAAGCACATGTGATGGCAATAACTTCTTCTAATATCACATAATATAGCA  
ATAGAAATAAAATTAAGAGTTTAGATTTTTGTGTAAAGGAGGTGAGATGTCACCTAATTT  
GTATGCTATTATGTAACTAGTCTAGGATATTGAAGCTGACTATACTCTGTTTTTAGGTCA  
TTATCTGTAGTTTACCATACTCCCTACTTGCTTCTTATCTACTATTTAACTCATTTTC  
CACATCCCTAATTTTGGTTTTCATGAATAATTTTTCTTCTGAATTAAGTACCTCTACT  
TACTATTATTAACCTTTATTTCTGACATATTTTATAACCTTCCATGGTCTCACTTGATTA  
AAAAATAAAAAATTCAGCTGGGTGCGGTGGCTCACACCTATAATCCAGCACTTTGGGAGG  
CCAAGGTGGGCGGATAATTTGAGGTGAGGAGTTGGAGACCAGCTGCCCAACGTGGTGAA  
ACCCCCCTCTCTACTAAAAAATCAAAAATTAGCTGGGCATGGTGGCAGGTGCTGTAAT  
CCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTGAACCTGGGAGGTGGAGGTTGCA  
GTGAGCTGAGATTGCACCTGTCACCTCAGCTGGGTGACAAGAGCGAAACAATGTCCTTGA  
AAAAAATAAAAAATAAAAAATCTCAACAACAGGGTTATTATTTTCCATTTTTGTTTT  
CCCTTATGAGTTTAAATAGTTTATGATTATAAACCTGAAAGCTTGAATACCTATGCTATC  
TTTTGTTTTCTTATGTTTATCAAGTTATTCCTTTAAACATTTTCTAAACTGTAAGAATAA  
TGTGAGGCTGGGCTCAATGGCTTATGCTGTAAATCCAGTGCTTTGGGAGGCCAAGGTGG  
GAGGACCACTTGAGGCCACGAGTTCAAGATTAGCCTGGCTAGGCAACATAGCAAGACCCT  
ATCTCTATAAAAAAATTAAAAAAATTAGCTGGGCATGGTAGCAAAATGCTTGTAGTCCCAG  
CTACTCAGCAGACTGAGGTAGGAGGAATGCTTGAGACCAGGAATTTGAGTGACCTATGAT  
TATGCACCTCCAGCCCGGGCAATAGCAAGACCCTATCTTTAAAGAAGAAGATGTAGTAA  
TATACATATTCATTTAATACTATTTTACCATTGAAAGTAAAAAATGAGTTTTCCTTTT  
CCCATGCCCCCTCTCAGAATGGGGATCTCAGTAGACCTTTAGGATTGGAAGAATGAGATC  
ATTATATTTTTCTGCAATTATTACCCACAAAATATTTAGATACCTTTCCATGTATTAC  
AAACAATGTGCATTTAACATGTCTCTCTCTTTCTCTCTCTCTGTGTGCGTCTTCATGA  
TCCTCTGTTGAGCCCTGCCAGTAAGACACTATCTCCTGAAGAATCACTGATAGGAACAG  
AAAGTGGAAGTGGCTAGGCCAGGAGTCCTTAGCTTCTTAGGGGGCAGGAGCTGCTTTGTGC  
TTTTCTCAGAATCAGATATATATGTGGACTGAAACATTTAAAAACAGAATAGCCAAGGGT  
CTATACGTTTAAAACTTATATAGATGGGCTACATGCTCTCTATTACTAATTTCCCAT  
ACAATACAGAGAGTGCCATGCTTTTTTAACTTGTTTTGTAGCAGACAGTAACTCTGTGTTA  
TGCAATGTTTTTTGATGAGAATAGGCTACTCATGAGAAATCTGTAAACCTAACACTAGTCC  
CTTGCACTACTCTAAATTGTTGCTAGAATCTTAAATTTTAGCACCAGACGGACCTTAGAA  
ATCATTAACTTTGGTGCTTTGTTCTACAATACAAGGAGATGGAATATTTTACCAGGATT  
GCTTAGCAGGTTACAGTTCTGCGCTCTGAGTACCAGCACTTCCCTGTGGGCAACATCAA  
CTTCTGATTTTCAAGTCTTAATTAGTACTCTGAAGAATCCTACTTGTTTTTAACTCCCA  
TTTTGCTTTGAAGTGACTTTACCTGATTTTTTTAGATCCCTTATTGCAAGCAATGCCACTAA  
GAACTAGTCTCTAGCTTCTTGGTGGGCAGGAGCTGCTTGTGCTTGCTCAGAACTCATC  
CTTTTCAGTAGAGGAGATATTGAAGAGAAATCTACTGAGGAGTCTGGGGGTGAGGCATC  
AGGGAATCTGCTCCAGTCCACAAAAGCAGAGAGGAAGGGTGGTTACCTAGAGTATTT  
AACATGCAGAGGCTTTGGATTTTACTCCTTTAATCCTTGAAATGCCTATGGAAGGGGAA  
AGGAAGTAAGATGGTGAAGTCCAGCTTATAGACATACTAGTGTTACATATATTTAACTAT  
AATAGGAGGGTATTATTAGTTTACTTTAACTTTCAACTGTGAAGGATTATACCTCTCAAT  
ATTTGTCTCCAGTGTCTATTTCAAGTGATTTTTCACTTTTCTTGAAGCAGCATGTCTGTT  
GCAAACTTCTAGAAATAATGAGAATATTTATATATTAGATCAAGCCATAACTTGATGAT  
ATAGTCACTTTCTCTTATATTTTTTACTTACATTTTACATTTTAACTTAACTTCTTCT  
TTTGAAAAACATGTCACTGCTGAGATGTTTTTCTTCACTCTGTAATTAAGTTATGAAAC

**FIGURE 14.3**



22/64

GTTTTTCCTAAATGCTGAGTATATCAAGTCTTGGCTAAGAATAAGTAATAAATATTGTC  
 CACATGAAAGACTACACATATAGCCAGGTGCAGTGGCTTGACCTGTTTTCCAGCTACC  
 CAGGAGGCTGAGGCAGGAGGATTGCTTGAGCCCAGGGTTCCAGGCTGCAGTGAACATG  
 ATTGTACCACCTACTCCAGAATGGGTGACAGAGCCAGGCCCATCTCTCAAAACAGAAA  
 AGAAAGATTACATAGACTACATATACACCCCATCCAAACATACACACACATCTACTTA  
 ACCTAAAATGGTAAGAAGATAACTTCTTATTTTCTAATATATGACACAGAAAAGTTTTT  
 TAAAGTAGTTTTAAATTTTTAATTTTTCTAGGTATTTCTCAAGCCATGTTCCCATGTGG  
TATCTTGTCAACAAGTTGAGGTGGAACCCCTCTCAGCAGATGATGGGAGATACTGGTAA  
AGAAAACCAATAAGAACTATCTCATTTAAGGTTAAATTACTTCACAATATCAATGTCTT  
 TAGCTTTCTCTAAGCTTTATTATATATTCTGAGTTGGTTTTGAATTATAAGAATGAATTG  
 GGGCCAGGCACAGTAGCTCATGCCTATAGTCCAGCACTTTGGGAGGCCAAGGCAGGTGG  
 ATTGCTTGAGTCCAGGAGTTCAAGACCAGGCTGGGCAACATGGTGAAACCCCGTATCTAC  
 TAAAAATACAAAATTAGCCAGGCATGGTAGTGCATGCCATTAGTCCCAGTCACTTGGGA  
 GGCTGAGGCAGGAGAATCGCTTGAGCCCGTAAAGTCAAGGCTGCAGTGAGTCAGGATCTT  
 GCCATTGTACTCCAGTCTCGAAAAACAGAGTGAGACCTTGCTCTCAAAATAAAAAAGAAATGA  
 ATTGATAGAGATCTAATGTACAACCTGACAACCTATAGGTAATAAAATTGTATTGGGGATT  
 CATGTTAAATGAGTAGATTTTAACTACTCTTACCACAAAAACAAAAAGTGGGTAACGT  
 GAGATGATGTATATGTTAATTTACTTCACTATAGTAACCATTATACTATCTATATGTAGC  
 TCATAACACCATGTCTGTATATTAAATATGCACATTAAAAATTGTTTTTTAAAAAAGA  
 ATTGAGATTTTTTTTAACTAGATATGGAGTGGACAAAATGTAAAGTGAATTGATCTTTTC  
 GCTGTGGTTCTAGGAGCTGCATGCTGTTCCCTTGAACAACATCTTCTAGATCAAATT  
CGAATAGTTTTTCCAAAAGCCATTTTTCTGTTTGGGTTGATCAACAAACGTACATATTT  
ATCCAAATTGGTAGGTGCTATTGTAATATTGCTGTCTATTTACACTATAGCATTGAG  
TCCAAAGTAGAAATGAATGTGCACTAATGAGCTTTATTTCTACACAGTTGCACTAATAC  
CAGCTGCCTCTTATGGAAGGCTGGAAGTACACCAAACTCCTTATTCAGCCAAAGACAC  
GCCGAGCCAAAGAGAAATACATTTTCAAAAGCTGATGCTGAATATAAAAACTTCATAGTT  
ATGGAAGAGACCAGAAAGGAATGATGAAAGAACTTCAAAACCAAGCAACTTCAGTCAAATA  
CTGTGGGAATCACTGAATCTAATGAAAACGAGTCAGAGATTCCAGTTGACTCATCATCAG  
TAGCAAGTTTATGGAATGATAGGAAGCATTTTTTCTTTCAATCTGAGAAGAAACAAG  
AGACATCTTGGGGTTTTAACTGAAATCAATGCATTCAAAAATATGCAGTCAAAGGTTGTTT  
CTCTAGACAATATTTTCTAGAGTATGCAAACTCTCAACCTCCTAGTATATATAACGCGTCAG  
CAACCTCTGTTTTTCTATAAAGCTGTGCCATTATGTTATTTCCATGGGACCAGGAATATT  
TTGATGTAGAGCCAGCTTTACTGTGACATATGGAAGCTAGTTAAGCTACTTTCTCCAA  
AGCAACAGCAAAGTAAAAACAAAACAAAATGTGTTATCACCTGAAAAAGAGAAGCAGATGT  
CAGAGCCACTAGATCAAAAAAAATTAGGTCAGATCATAATGAAGAAGATGAGAAGGCCT  
GTGTGCTACAAGTAGTCTGGAATGGACTTGAAGAAATGAACAATGCCATCAAAATATACCA  
AAAATGTAGAAGTTCTCCATCTTGGGAAAAGTCTGGGTTAGTATAAATTTATAACTTGGG  
 AGAAATTTTATGTGGCTTAAACATCCCCAAATTATGAATTAGAATAGTATTTCATATATA  
 AATTGAAAATCAATTAAGAAAGAACACAGTGCCTAAAGGCACTTGGGGGACACATTTACG  
 CTTTGCAGTAAAGTCTTTGTTTGATAAAGATTGTATGTTTTCTGGCCAAGTAAGCTTGA  
 ATAGGTACAAGCTTAGATAGGTTCAAGCCAGAGAGGTCAAAATTACTTGCCCTGAGATTGC  
 ATAGCTAGTGTTACAACCTAGGATTCAAACCCAGGCAGATTGACTTGGGGGTTTCATCAGGA  
 TGGAGTGCCCTACAAAGCCTCCCATCTTTAATGCTTGCAAGATTGTTCCCCAGTTACCGA  
 AAGCAACTTGTTAATATAGGGAAAAGGCCAGTGTAGGGAGAGATCCATGGCATGAGGT  
 AACCTTCTGCTGCATGTGGTGGCACCTGGATTGGAATGCATCCAGGAGCTGCTTACCCT  
 GCCGGTGTCTGCTCTTTAATTTGTGTATAACGGAGAGGAAGTAGACAGGGCAACTAGTGC  
 TCCAGCCCTCATCTGGCCACAAATATTAATGTACCTTTATATGACATAAGTCACTAG  
 TCCATTTATTGGAACCTAAATTTGAACCACTGTAAGTAAGACTTCATAGTGATAAAGAG  
 AGGAACCTTGTAGGAAAGAGAAATAAATAGAAAGAGAAGGTTGTCTCCTTTTGTAGATTT  
 TTTTTTTTTCTCAACAGTTTTTACCTGTGACCTTTATACAAATAACTGACAAAGCATTAA  
 TCTCTTTGGCCTACATCTTTTCTTTTCTATTTTTTTTTTCCACAAGATGGAGTTTCACT  
 CTCTTGCCCAAGCTGGAGTGCAGTGGCATGATCTGGCTCACTGCAACCTCCGCCTCCCA

FIGURE 14.4

23/64

CGTTCAAGTGGTTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGCATGCACCAC  
CACGCCTGGCTAATTTTTTGTATTTTTAGTAGAACTGGGTTTCACCATGTTAGCCAGCC  
TGGTCTGGAACCTCTGACCTCAGGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGAT  
TACAGGCATGAGCCACTGCTCCTGGCCGGCTACATCATTTTTCTAAAGCTCCAGACCATT  
CTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTC  
TTCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTTCTTTCTTTTGTAG  
TTAGAAGCTTGCTTTGTGCCCCAGGCTGGAGTGCAAGTGCCACCACCTCCACTCACTACAA  
CCTCCACCTCCCAGGTTCAAATGATTCTCCTGCCTCAGCCTTCAGAGTAGCTGGGACTAC  
AAGTGTGCGCCACCACCTCTGGCTAATTTTTGTATTTTTAGTAGGGACGAGGTTTCACCA  
TGTTGGCCAGGCTAGTCTTGAACCTCTGGTCTCAAGTGATCCGCCTGCCTCAGTCTCCCA  
AGGTGCTGGGATTACAGGCGTGAGCCACTGTGCCTGGCCTCAGATCATTTATTTCTGTTA  
GCTTTAAACTGTCCGTTCAGGAGATCCCACTGCATCCTCAAATTCAAAATATCTAACACT  
GAGCTTATGATTAGCTGGTCTGTCTATTAGATGGGAATATCCTTTTATTTCTTTGAAAT  
TATATGGTGAGAACAGGAGAAAGTGCTGATGGTAAAGTCTGTGATTAAAGATAGCAATAA  
GGACTCCGCCCTTCCCCTCCACTCCAGGTTGAAGAGCCATGGACAATGAGAAGTCACAG  
TAGGTGAAATCAGGTACTAAAATGGACTTGGCTTGAGAGATCAAAATTGATCACTTGGTG  
ATACAACTAACAAATTCATGTAACTTGAACCTTTATTACCCTGTGAAGCATGGTGATTA  
AAAAAAAACAACAACAACAGGAACTTGATTGTTAAATTCCTTTAAGTCAGAAATATG  
TACCTTAGAGTTTTTATTATGCTTTTGTCTACCATTAAATATGTCTGCACCTGCTCTTTA  
GAAGTTAATAGAGAGTAAAGTCGTCTTTATGTCTTTCAGTGCTTACTTATATTTGGGAAG  
TTGAGAAAAATTTTAAACATCATTATTGATATATATATATATATATATATATATATAT  
ATATATATATATATATATAGATAAATTTTTTTTTTTCTTGAGACGGAGTCTCACT  
CTGTGCGCCAGGCCGGAGTGTGGTGGCGATCTCCACTCAATGCAAGCTCTGCCTCCCAGG  
TTCAAGCGATTCTCTGCCTCAGCCTCCCAGTAGCTAGGATACAGGCTCCCACCACCAC  
GCCTGGCTAATTTTTGTAGTTTTAGTAGAGACGAGGTTTCACCATATTGGCCACGTGGT  
CTCAAACCTCTGACCTTGTGATCCGCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGG  
CGTGAGCCACTGCGCCCGGCTGAGGTAAAAATTTAAAGTGACAAATTCAGTCATTTTTAGT  
ATATTTATACTAGTTGTACAGCCATCACCACAATCTAAGTTTAGAACATTTTCATTAGGG  
GGTGGGAGAAATTTTACTCTGCTTTTATAGATTAAAGTTCTGTCTGGATCTAATCATTTAA  
TCAGACAATCAGGCAGATTGTCTGTGATTAGTTTTGGCCATTCCAGCTTCTTCATTGGTT  
GTTAACTTTCACAAATAAAGGCTGCTCAAAGATTAGAAATAACATTTAATTTGAATGTAA  
ATGTGCCATAGTTTAAAAGATGGGTTTGGTGAATACAGTCAAATACATACATTTAAAGCT  
CTAATTCTGAAGATTATGTAAAGAAAAGGAAAGAAATGTAGGGAGAGGATTGAAATGTTT  
ATGGTATAACAATATCTGAACATCCATCTGGTCACACCGTTGGTATTGAAATGTTTGTCT  
CTCCTCAAATTCATATGTGAAATCCCAACTCCCAAGGTGATCGTATTAGGAGGTGTGGT  
CTTTGGGAAGTGATTAGGTCATGAAGGTGAAGCCTTCATGAATGGGATTCTGTCTCTTAT  
AAAAGAGAAGTGTGAGAAATAAGTTTCTGTCTGTTGTTAGCCACCCAGTTTAGGATATTT  
TGATATAGCAGCCTGCATGGACTGAGACAACATAGAGTTATTATGATAGCTTCTGTATT  
TCACCTAAATTCATAGAAGCTAATATATCAATATTTATGCTATGAAATATTTCTTAACCA  
AGCTTTGAATATATTTATATTTTGTATTTTAAATTTTCAAGATTCCAGATGACCTGAG  
GAAGAGACTAAATATAGAAATGCATGCCGTAGTCAGGATAACTCCAGTGGAAGTTACCCC  
TAAATTTCAAGATCTCTAAAGTTACAACCTAGAGAGAATTTAGTGAGTTCAAATATATA  
TGTTATCATCAAAATTTCTTTACAGTTTTTGAAGATTCTAGTTGCTTTAGCTAAGTAAT  
AAGAATGTTGTATTCTTTTGTATACAAATCTTTTTTATTGTGTTAAACTATATATAAC  
ATAAAATATGCCATGTTTCGCCATTTTAAAGTGATAATTCAAAGGCATTAATTACATTCA  
TAATATTGTACAACCATCACCCTATCTATATCCAGAACTTTTCCATCACCCTCAAAGAGA  
AACTTGGTACCATTAACAATAATTTCCCGTCCACTCCTTTCCCCAGTCCCTGGTAATC  
TCTAATGTATATTGTGTCTCTATGAATTTACTTATTCTAGATATTTTATATATAAGTAGA  
AGTATGCATTTGTCTTATGTATCTGACTTATTTCAATTAACATAATGTTTCAAGGCTCA  
TCTGTGTTGTATGTATCAGAATGTTATTCTTTTCTGAGCTGAATACTATTCCATTGACT  
GCATATACCACATTTGTTTATCCATTCTGTGATGGACACTGGGTTGTTTCCACAT

FIGURE 14.5

24/64

TTTTGGCTGCTGTGAATAATGCTACAGTGAACATTGGTGTACAAGTATCTGTTTGAGTTC  
 CTCTTTTCAGCTCCTTTGGGATATACCTAGGAATTATGTTTAACTTTTGAAGCTGAG  
 AAATCTTTAATAAATGATAACACAAATACTTATATTTGCCAATGCAAATATGAATATTTT  
 TGGCTTTTAAGAGATTGATCATTTTGCCACGTGGTGTAAATAAAAAAATTGTCCCATG  
 TTGTTTCAGTATTAATATTGTAGCCTAAAAGAGTGCTAGACTGTTTTACTTTTACTCAG  
 TTAATTCCTTTGGATACTGGTAGAGTCAGGAAATGAGATATTGAACTTAAAGATCTTTGCA  
 GGTGGGGTCCAGTGGCTCACACCTGTAATCCTAGCACTTTGGGAAGCTGAGGTGGGAGGA  
 TTGCTTGAGGCTCAAGAGTTTGGAGATAGCCTGGGCAACATAGCAAGACCCCATCTCTACA  
 AAAAAATTAAAAAATAAAGCCAGGCGTGGTAGCTCACGCCTGTTATCCCAACACTT  
 CGGGAGGCTGAGATGGGTGGATCACTTGAGGTGAGGAGTTGGAGACCAGCCTGGCCAACA  
 TGGTGAAACCCCATCTCTACTAAAAATACCAAAATTTATCGGGCGTGGTGCTAATCCTGT  
 AATCTCAGCTACTCAGGAGGCTGAGGCAGGAGAACCACTTGAAGTGAAGGAGGTGGAAGTT  
 GCAGTGAGCCTAGATCTCACCCTGCACTCCAGCCTGGGTAACAGAGCGAGACTCTATTT  
 CAAAAAAGTAAAAATAAAAAATTAGACACATGTGGTGGCACATGCCTGTAGTCTTAGCTA  
 CTCAGGAGGCTGACTGAAGTGGGAGGATCTCTTGAGCCAGGAGTTCCACACTGCAGTGA  
 GCTATGATTGTGCCACTGCACTCCAGCCTAGGCAATATCTCAAAAAAATTTTTTAAAT  
 AGATTATTAGGCCAGAGCTGGTGGCTCATGCCAGTAATCCAGCACTTTGGAAGGCCAAG  
 GCAGGCGGATCACCTGAGGCCAGGAGTTGAGACCAGCCTGGCCAACATGGTGAAACCCC  
 ATGTCTACCAAAAAATACAAAAATTAGCTGCAATGTCTATAATCCCAGCTACTTGGGAGCC  
 TGAGGCAAGCGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGTGGAGACTGCGCC  
 ACTGCACTCCAGCCTGGGCGATACAGCGAGATTCTGTCTCAAAGAAAAGGAATTTGTTT  
 TCCTGTCTTTATCGTAGAGGAGGAAAAGGAGAAATGGGGTTGGAATGGTTATTGAGTGAG  
 CCACATTATGGTAGATGTATCACTGGGCATAGAGAAAAGGAGCATTTAAACTTTTCCGC  
 CTAACAGATGTTTTCTTCAGGCTACACTGCACTCATTGTGCTAACTGTAATGTCAAATCCC  
 AGACCTGTGCCTATAGAACATGAACATCCTTCATTGGATTGTTTGGTCAGGCTTACACT  
 TTATTAGGAAGATCAGATGTTAAATTAAGGGTGTAAAGTTAAGTTCAGATATGAGGATA  
 ATTCACTACTATTCCTTTTCTGGCAGCCTAAAGACATAAGTGAAGAAGACATAAAACT  
GTATTTTATTATGCTGCTACAGCAGTCTACTACCACCATGCTTCCTTGGTAAATACAGAG  
GAAGAAATTTATTAAGCTGGAACTAAAGATGGTGAGTACATTTGTTATTTTGACTTTTTT  
 TTCTATTAAATAGTTGTACATTTTTTAATTGTTCTTGCAACCTGTCATACCTGTGAACAG  
 TATGTGAATAGTGAAATATAATTATGATAATTAAACAGTAGTTTTTATGTATTGAAAAAT  
 ATCTTTGGCCGGGTGCAGTGGCTCATGCCTGTAATCCAGCACTTTGGGAGGCCGAGGCA  
 GGCGGATCACTTGAGGCCAGGAGTTCGAGAGCAGCCTGCCAACATGGCGCAACCCATCT  
 ATACAAAAAATACAAAAATTAGCCTGACATAGTGGTGTATGCCTGTAGTCCCAGCTACT  
 TGGGAGGCTGAGGCAGAAGGATCACTTGAGCCCAGGAGGTCTGTGTTCTGCCACTGCAC  
 TCCAGCCTGGGCAGCAGAGTGAGACCCTGTTGGGGGAAAAAAGTCTTTAACTT  
 AAATAAATTTGACATTTAAATCTTAAATTATTCATCTCTGTTTCAGTACTAACTCTGC  
 ATTTATTACTTTCTTTTAAATAGGACTGAAGGAATTTTCTCTGAGTATAGTTCATTCTTG  
GGAAAAAGAAAAAGATAAAAAATTTTTCTGTTGAGTCCCAATTTGCTGCAGAAGACTAC  
AATACAAGTAATAGCATGTTATTGAATATTTAATAAAATACTATTGTTACATATGATTG  
 ATAATAAAGTATGAAGTTCCTTGTAACACCTTGCAATGTGAAGTGTATTAACAACTGCT  
 AAGAGTAAGGAATAACTTGATTTAAATATTTTATTCTGTAATCTCTTAAATTATCTGT  
 ACAAAATTATTGACTTAACCTAAATTTAAAAATGAATGCCTTAGCACAATTAAGTCCAAG  
 AATAGAGTTGATCATGTTAACTGGTAAATGGATCATGATTTAAATCTTCTAGGATTGA  
 AACAAATGAAAACGTAGTTTTTAAGGGTTTGATTTTTTAAATTCCTATTTTTTACATGCAAT  
 TTTACTGCACAACCCATCTTATTTTGACAGTCTTAAATTCGCAACTCTTCAGAAATATT  
 ATCAGATCACTTTCTTTGCTTCCATAAGTTTTTTATTATTATATTATTTTTTTTTT  
 TTTAAAGACGGTGTCTCACTTTGTGCGCCAGGCTGGAGTGCAAGTGGCATGATCATGGCT  
 CACTGCAGCCTCGACCTCCAGGCTCAGGTGATTCTCCACCTCAGCCTCCCAAGTAGCT  
 GGGACCACAGGCGAATGCCATGATGCCTGGCTAATTTTTGTATGTTTGTAGAGATAGGG  
 TTTCAACATGTTGCCGAATTTGTCTGAACTCCTGGGTCAAGCAGTTGTTCTGCCTTG  
 CCCACCCAAAGTTGTGGGATTACAAGTGTGAGCCACTGCGCCAGCTATTCTAGAAGTAT

FIGURE 14.6

26/64

CTGATGTCCTTGTGTCAGGACTTAGGAATGGAGCTCTTTTACTCACAGGAGGAAAGGTA  
AGTGGTTAAGGTGTGTTTCATTTTCTGTAACATTTAATACTTTTCATTATCTTTCTTT  
GGGTTTTGACCATCTATTATATAGGGTGGGTTTTGACCATCTATTATATAGGGTTTATAC  
GACATATGGAAAGCATTCAATTTATCTACTAATATTTCTGTGTGCTGCTTTTAGGTGTTG  
GGGGAGTGATGACGAATAAGACTGATGTTCTCCATGCCCTTTTTCTGTGTCAGTTGATAC  
AATTATATGGTTTTCTTTTTTAGGCTATTAGGTGTTGATAGGGTTGAGTAACTTACAAA  
TGTTGAACCAAGCCTTGATACCTGTGATAAATACCACGTAGTTGTGGTGATCATCTTTT  
CTACATTGCTGAGTTTTATCTGCTAATGTTCTGTTGAGCTTTTGTCCATTTAAGTTTGAA  
AGTGATTAGTTTGAGTTTTCTGTTTTTGTGTTGTCTTTGTCTGGTTTTGCTATCCGTGT  
AAATCTGGCCTCATAAAATGAGATGGGAAGTATTCTCCTCTTCTTTGTTTTTTTGGGA  
AGAGGTTGTATAAAATGAGGCTGAATCTTGGTGGTTGCCCAATGACAGGAACATTTTC  
TGTGACTGAATATATTGGGAATTCCTATAAAGCAATTATTTTCTAGGGAAGTGGAAAATC  
AACTTTAGCCAAAGCAATCTGTAAAGAAGCATTGTGACAACTGGATGCCCATGTGGAGAG  
AGTTGACTGTAAAGCTTTACGAGGTATGAGTATGGTAACACTCTATATAAATCCCTTTTT  
CATTAGAAAGACAGGAATGTTATACATAATGCTGTCAATCTAATAAATACACATATCATC  
TAGTCTTTAACTTTTCTGTTTATCATTTAGTCATTAAAATTTCTTTGGCTTTCTAATGTT  
TTTGATAAAATTTCTAAAATCTCCATATTTAATGGAGGCCTATTTTTTTTTCTAGCCAG  
AACTTTTTTGTAGACTACATTTCTGGAAGTGCTCACTGACACCACTCTGAAAAATTAGTAC  
TTAGAATATACTCTAATTGGTATAAATGATCTCTGAATTGCTATGGAAAATCGGGAGAAT  
GGTTGCTTCAGGGGAGAGAAAGTAGGAGGCTGTGGACAGCAATGAGGAGAATTACAGTTC  
ACCATATAACACTTTTGTACTTTTAAAGTCCTTAACATTTACATTATTATCTATTCAATT  
AAAAAATATTGGGAAGATTTTACTTTGAACAGTTAATTTTTCCCCCATGGGTACCGCTGT  
CATATAGTTCCAACATCATGAACCTGTGTATTTCTGTTCTTTGTAAATTTAAACTTT  
GTAACCTACCAGGAAGTTTGAAGCCAAATTTGTGTTTCAAATATAGCAACTCCAGGATCT  
CTAGGCAGATGCATTGCAATTGATTTTAAATGAATCTTGATCCCTTACTCTCACTTATG  
TTTTCCCATCTACTTTTTTTATTTTGTGTAAGCCATCTAAAATTTCTCAATGGGATG  
AACTGGGTATAAATGAATACATGCATACAGGAATTATAGTAGCATATTCCTTTTCTTTT  
TTCTTTTTTTTTTTTTTTGAGACAGAGTCTTGCTCTGTAGCCAGGCTGGAGTGCAGTGG  
TGCGATCTCGGCTCACTATAGCCTCCACCTCCCAGGTTCAAGCAATTCTCGTGCCCTCAAC  
CTCCCGAGTAATTGGGACTACAGGTGCATGCCACCACCTGGCTAATTTTTGTATTTTTT  
TAGTAGAGATGGGGTTTACCATGTTGGCCAGGCTGATCTCAAACCTCTGACCTCAAAGT  
GATCTGCCTGCCTTGCTTTCCCAAAGTGTGGGATTACTAGCATAAGCCACTGCACCTGG  
CCTCCTTTTCTGAGTTTTATAAAATTTGATACTTTACTGCACGCTTTGAGACTGTATTAA  
TTGAACCATGTTGATGAACAAGTTTTTGTGATGGGTATATTAATAAAATATAGATCAAAT  
TTTATAGTTAAATCAATATCGAGCTTTTCTAGTGCTTTCAAAGGACAACCTGAATTTT  
CCCGACTGAAATGATACTGAAACCATTTCAATCTTCTGTATTAAAGGAAAAAGGCTTG  
AAAAATACAAAAACCCTAGAGGTGGCTTTCTCAGAGGCAGTGTGGATGCAGCCATCTG  
TTGCTCTGCTGGATGACCTTGACCTCATGCTGGACTGCCTGCTGTCCCGGAACATGAGC  
ACAGTCTGTATGCGGTGCAGAGCCAGCGCTTGCTCATGGTAAATGCATCCACCACTGGC  
TTAAGGTCTTGTTCTTTTGTGTCAGTCAGCATTTTTTAGTCTTAACAATAAATCTACTCTCTT  
CAGAGAATAATATATGTGTTATGTTAAGTGTGTTGAGGCCCTGATGGCATTCTAC  
AGTTGTCTATAGACTGTAATAGCAAAATTTGGTAGAGTAAAAACAGTGTGAAAATTTCTGC  
AACTTCATGGTTAGTCCCTTTAGGGTTTTTCATTCTCCCTTACTTATTGTTTTAATTACAG  
ATTTACTCTTTGTTTCATTTGACAAATATTTGTCAAATGCTTGTGCACAGTCTGTATTCT  
CAAATTTAGGAGAAAAAGAAGGGTGAACAGTATTAGCGCAGAACGATACTAATAATGAT  
GGCTACTGTGATGAGTAGCCAGCCCTTTCTTGCTTTCTTGGATTGCTTTGTATTCTAC  
ATGAAGATATTCCTGGGCTTTACAGGTCAATAAATGGAAATTCAGAGAGATTAATTTGA  
CCAGGGTGACCAACAAGGAGATGACAGCATACTATGCGAGAAGTATACACAGAGTAGT  
GTAGGAGCATATAACCTAACTGGGGGTGAGGTGGGATAAGGAGTTATCAGGGAAGGCTT  
TTTGGAGGAGTTGACAACTGAGCCGAGTTTTGATGGAAGAGTAGAAAATTAGCATGAACCA  
ATTTTCATGCTAATAAAGAAGCAAAGGAAGCGTGGTCTACAGGCAAAGCACAGAGGTACA  
GGAAGTAATGATATGTTGGGAATACCTGTTGACTGGAGCTTAGAGTGCAAGGAGAGGA

FIGURE 14.8

27/64

GTGCTAGGGAGGTGAGGTTGGAGGGTTTGGCAGCATTGACTTGCTTCAAGGTTCTTAAGA  
 GCTGAAATAGATATAAAATGCAACTAAGAGTGGCTTGGATTATTATTACCTAGTGTGTTA  
 ATCTCAAATTTTGAAATCTATAGCATCTATAGGACTGGTGTACTAATCTTACACTCGAT  
 CTGTTACTGTTCTTATACCTAGATCTATTAGTCCAGTGTTTAAGGGAGTGGTGCAGATTTC  
 TAGGTGAGGACAGGACTCAGATGTACATTATTAATGCCTATTTTCAGTTCTGACCTTCTCA  
 TATGAAACCTTATAAGACCTGGGGTAGGAAGAGATTGTTCTGGAAGTCATAGGAATATGA  
 ACTGTATTTTGTTTAACAACAATACAGTATGGAAATTTATCACCTTCCAGAATATTTA  
 TTTGAGAGACAAATTTTATCATTCTGTTTATTTTTCATAAGATCCACGAGTAGGGAAC  
 CTCCTAGACATTGCTCTGAGTATATGGTCTGAGTTTGCAGTACCTCTTGTGTCTCCATT  
 AGATTTATTAGGTCCTCAATAGATAAATCAGGGAATAACTAGATGGATTCAATTTTTTAAA  
 GACATGAAAGAGCGCATACCATACATACTGCACCTTAAAGGTCAACCTTAGAGTATCATT  
 TTTTAAATGAATGTATAATTTTAAATTTTATGTTTACTTTTCTTAAGCTTTTGCAGTAT  
 ATTGCTTAATTCAGCTTTGAATGATATGATAAAAGAGTTTATCTCCATGGGAAGTTTGG  
TTGCACGTGATTGCCACAAGTCAGTCTCAGCAATCTCTACATCCTTACTTGTCTGCTC  
AAGGAGTTCACATATTTTCAGTGCCTCAACACATTTCAGCTCCTAATCAGTAATACACT  
 ACTTGTAAAGGATTATTGAATTATGTCCCTTTTATAGAAATATTTTTCAATTTTATTAGT  
 AATTCTGCTGGCTTTAAATTTATGCTTCTCTTAATGATTTTAAAGGATATGAAGTCAACATT  
 TGGTGCATATTGTGTAGAGCATATAAATTATAATTTATAGCCACCTGAAATGTTAGTATG  
 CGCTTTCCAAGAAAATGACTTTTTTGAATGGTATTCTTTGAATGAGAAAGAACAGAG  
 AGAAATAGATAGATGGCTTTTAAACACTTCATTAATTAACCTTTTTTTTCCACCATCAC  
 ATAATGGCATTAGTCCCTTTTGGGAACCTCATGAGGGTTTATGTTGTTAGTGTGAAAG  
 AAATATGTTCCAGGACTGGCAAACATATTCTAAATCTTTAAATTTTTCACCTAGCATCT  
 ACCCTAAATATTAGACCTGTGCTAGTTAACTGCTATTGAAGAACAAGGTATTATATC  
 TATTATTAAGGATAATAGAATGGTATTTGAGATATTGGTCATTGAATATGAATATGTTT  
 GAGAAATAAGTTTATAGGAACCAAAAAAATTTCTTAAAGGAACCATATATTACTAAAA  
 ATGCTTCTTATTGGAGAAAGAAATGACAATCATTTATTAATGTGATTTTTTCAAACTTT  
 ATTAAGATATAATTTAAGTACAACAACTCACATAAAGTGATCAATTTGATCAGTTTAA  
 CATATGTAGATGCCATGAAACCATCACCACAATTAAGGAAACAAACATTTTCATCACTCC  
 AGAAGTCTCCTAGCCCTTTTACTACCCATTCTCCCTGCTCCATCCCCAGACAACCTACC  
 AATTTGCTTTCTGTCACTATAGATTTGTCAACCTGATTTTCTCCAAATATACATTCAAAA  
 ATATACAGTTGAATAATAAATGGAAATTCGAATTTTGTGTTTTTTCTTTAGGAACAAAGA  
TGTGAAATTCGTGTAATGTAATAAAAAATAAATTTGGACTGTGATATAAACAAGTTCACC  
GATCTTGACCTGCAGCATGTAGCTAAAGAAACTGGCGGGTTTGTGGCTAGAGATTTTACA  
GTACTGTGGTTCGAGCCATACATTCTCGACTCTCTCGTCAGAGTATATCCACCAGAGAA  
AGTATGTTTTACTATTAAACCTGAACCTTGAATCTTCTTTCTATTGTGGAGAAATGTAA  
 TTGTAGTAAGACAAGAATTAAATATATTCCATTGTAGTATTTGAATAAGCAGTTATTTGA  
 GTAGAAAATTAGTGTTTCCAGCTAAGATGATGGCATATTTTGAAATTCATATAGTGAAT  
 ATAACTAGTAAAGAAGTTTGTGTTATTTTAAACAGAAATTAGTTTAAACAACATTGGAC  
TTCCAAAAGGCTCTCCGCGGATTTCTTCTGCGTCTTTGCGAAGTGTCACCTGCATAAA  
CCTAGAGACCTGGGTTGGGACAAGATTGGTGGGTTACATGAAGTTAGGCAGATACTCATG  
GATACTATCCAGTTACCTGCCAAGGTATGTTAAAAAAGAAAAAGTGAATACTTACTCC  
 CAGAAGAACCCTGTATTATTGGCTTTGGCTTTATGTGTGAGCTTGCCCAATCTCCGTGT  
 GAGTCAACAAGTGTCTTACTGAGTTACCAATAAATGTCTTAACACTATTTTAGGTACTTT  
 AACAAATTTTAAATTTTATTAATTAATTTTTTATTAGAATTGAGACCTCACTCTGTCTCT  
 AGGCTGGAGTACACTCACAGCTCACTGCAACCTCAAACTCCTGGGCTCAAGCAATCCTCC  
 TGGCTCAGCTTCCCGAGTAGCTAGAATAACAGGCATGAACCACCATGCCGGGCCAATCT  
 TTAATTTTCTTAGAGACGGAGTCTTGCTATGTTGCCAGGCAGACAGATTTTAAATGTGTA  
 TGATGCAGTCTTTGATGATAAGAACTTATAATGGAAAGCTGAGGTGATAGTTACAGTAA  
 ATACATTTTGTATGATAATTCTGTTTGTCTTAAATCATTCAAATGTAGTAAAGCAAGATG  
 AACTGTCTGCTGGGATTTGAGCAGAAATGGATAGGAATAAACTAGGAGGTAGAAGAGTTA  
 TCAAGGTTACAGGACTGATGGGTGAAGCTAGATTTCCAGACCCGGGATGTGAGTCTCTG  
 AAAAGCAGACTTGGCAGGCATAGACGAGGCAGATAGCAGGATAAAGGAGACAAATGTAGA

FIGURE 14.9

28/64

TTGTTCTTCAGAAGATCAGATGGTAGAGTCTAGGAGGTAGTGTGTTTTAATCAGAGATCT  
 GAGAGGCCAAAGATCATTGCATGAGATCAGGGACCCATGCAAAGGAGTGAGAAAAAACT  
 GGGTTAAGGAGCCTGCTGCATGGCAACTCCTGGGAACAGTGGCCACTGGGGCCTGGGACA  
 TGTGATTGCAGCCAGGACTGTTAAAACAGTGTGAGAGAACATGGGTATGGAAGTACT  
 AGCTAAGCAGGATCATGACCCCGATGCTGGGATGGGGCATCAAGCATTAGTACATGGAGAT  
 TCAGTACATCCAGATGCAGTACATGGAGACTATATGCGTAAGTCTGACTTTGGGCTTCT  
 TTCAGATTGGAGCAGAGGTAGAGGTGAGTGGGAATATTCTCAATAGAGGGAACATAATAG  
 GCATACCTAATAAAGGAGACCAGGATATTGCAGACAGTAGCCTCATGTTTGGCTCACCTG  
 TTCAAAAAGTCTCTTGTCTTGTAGCAGTGGTGCCTTAAAAGGTAACCTCAGAAGCAGTC  
 GATTATTTGTTTCCAGCCTGGAGACTCTTGGGATATTTTACTATCTTTGATTGAATAGATT  
 AAATGTACACAGCTCTCATAACTTGCCCCATGAAGCATATCCATGAAAGGCATACTT  
 GTTAAAAGATTGGTTTTGTACTTTTAAATGTAGTACTTTTAAATAAACAGGAAAAATAGA  
 AGTCTGATGCAGTTATATGCATTTTATATAGAATGTGTTCTTAATTGGAAAAAATTTGT  
 CGTAGTTCCCTTGAGTTCATTTACAGTTTTTAGTAGGAATTGTATTTTCTACTGTTGTAC  
 TTGCTGTTACTAAAGAAAGATGGTCGTGATTACCATCTGAATTTTTTTTCTATACATTGA  
 TCTTTAGCTGCTACTTAGTCATTTCTGTTTAGACTTGAGCTCTTTTTCATATTTTTTTTT  
 TTTGTTTCTCAGTATCCAGAATTATTGCAAACTTGCCCATACGACAAAGAACAGGAATA  
CTGTTGTATGGTCCGCCTGGAACAGGAAAAACCTTACTAGCTGGGGTAATTGCACGAGAG  
AGTAGAATGAATTTTATAAGTGTCAAGGTATGTTGTCTACTTATCTTCTTTTTTTATTTA  
 GGTAAAAATTAACATAAATGCAGTTAGCCATTTCAAAGTGTAATTCAGTGGCATTAGTG  
 CATTCAATGCTATGCAACCACCACCTCTCTCTAATTTCAAACCTTTTTTATTCCACTC  
 CTCCTCTTGCTTATCCCTGGCAACCATTCATCTGCTTTTTGTCTCTATGGATTGCGCTT  
 TTCTGTATATTTATATAAAACAAATCATGCAATATGTGACCTTTTTTGTCTGGCTTCTT  
 TCACTTATGTAATGTTTTCATGGTTCATCCAGGTAGTAGCATGTATCAGTACTTCATTCC  
 TTTGCTGACTGTAATAATGTTACCATACTTTGTTTATCCACTTATCAGTGGTGAACATTT  
 GAATTGTTTTCTACCTTTTGACTATTATGAATAATGTGCTGTAAATATTCATGCACAAAT  
 TTCTCCACGGATATGTTTTTCTTCTTGGGTATAAACTGAGGAGTAGAATCTTGGGT  
 CTTAGGGTAATTCTCTAACTTTCAAAGAACCACCAAAGTCTTTTACACCAACTGCAC  
 CATTCCCACTAGCAGTGTGGGGGTTCTGATTCTCCACATCTTTACCAACACCATTATG  
 TTTCTCAATTGTGGGCTAGTCTCACATTTGGAAGCTAGTGGGAGCAGCGATCCATCTAT  
 TAAAAGTTGTATGAAATTGAGTAATGAGCCACCTCTCTCTTGTAGGGCTTATTATGTTCT  
 TGCTTAAGGCAATCTTCATGCATTGTGAACAGAATTATACATAAATGCTCAGATAAAAGG  
 GCAAACCATTTCTTAAAGGGAGTAGACAACCTAGAGGCAGGAGACCATACTGAGGCAGGAAG  
 CTGGGGTTTTATGGTTCTGTTACTTTTGACTATATCTCACCATTGCTTTTGTCAAAGTG  
 AGACTAGGTCTAAGTTTTTTTTCAGGTATAAGGTGAGTGTGGTAATTAAGGGGCATGCTAG  
 CAGATCATTTTGGGTAATGCTTACAGTCCACCCTGGTGTGTCATTGTGGTTCGAGATC  
 CAGTATCTTAGCTGTGTAATTTTACAGATCAGCAATATTAGTTTAAACAAAGGGCAATTAG  
 ATTCCAAGACAAAGGAATCGTGTATTATTCTAGCCTTATTCAAACCTTGATTATATAATCA  
 GTTTAGTAATTTATTTATTTGTTTCTGTATTATTTTATTTTCTTTGAGATGGAGTCTCA  
 CTCTATTGGCCAGGCTGGAGTGTAGTGATGCAATCTTGGCTTACTGCAACCTCTGCCTCC  
 TGGGTTCAAGCTATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCTAATTTT  
 TGTATTTTGTAGTAGAGATGGGGTTTACCATTGTTGGCCAGGCTGGTCTTGAACCTCTGAC  
 CTCGAGTGTATGCCCCGCTTGGCCTCCCAAAGTTCTGGGATTACAGACGTGAGCTACCG  
 TGCCCAAGCTCAGTTTATGTAATGTATAACTGGGTTTTTACCAGTTGTAAATTACTCTTTG  
 TCGTGTTTTTTTGAGAACTGGCAATGACGGAGAACTAAAAGTGCCAGGCTGTTGCCTTG  
 TTCTGTATTTTGGCCTTAGTTTTTTTTTTTTTTTTTTTTTCTCTGAGACTGAGTCTTG  
 TTTGTTTACCAGGCTAGAGTGGAGTGGCATGATCTCGGCTCACTGCAACCTCTGCCTCCT  
 GGGTTCAAGTGATTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCGCTGCCACC  
 GCACCCGGTGAATTTTGTATTTTGTAGTAGAGACGGGATTTTACCATTGTGGCCAGGCTG  
 GCCTCGACCTCCTGACCTCATGATCCACCAGCTTCCGGCTCCCAAAGTGTGGGATTACA  
 GGGGAGAACCACCGTCCCGGCTTGCCTTAGTTATTTCTTGTTCCTCCTCTAGTCCTA  
 TAGTTCTCTGACTGTATTGAGGAAATGTAATTAATATTATTATGTTAATAGATATTTAT

FIGURE 14.10

29/64

GTGGTTGAATATTAGAAATTCCTTATTTTGGTCACATATCCTGATCAGTAGTTGGTCTTC  
 TGGAGATAGTGATTTTTCTACTAGAGATGACTTTAGGACCTATTACAGGTTTTTTTAAAGAT  
 CCCAATTTAAGGAAAGACTATTCTCATTATTGATTTTGCTATATGCAGGGAAATTTATTT  
 CGAAAGGTTTTTCAGTTGGCTTTTAGGGAAGATTATATATTCTCTTTTTTTTTTTTGGC  
 CTTTTCCACATGTTCTAAAAATGATATATTCTTTAACTCCTATGAAAAATACATTGTTTC  
 AGTAATTGAAGATGCTGATTAAAGTCATATCTCTACACATTTTTTAAAAATTTGAGATAGA  
 TGGGACTTTGTCCCTTCTTACACCATTCACTTATTCACTTGGAAAACTATTATCCAATA  
 CTTATGTGGCAGACACTGTTCTGGCACAAGGATTTCAGCAGTGAACAAAACCTGCCTTTT  
 TGGAGTTTACATTCTACTAGTGGAAAGCGACAACAAGCAGATAGACACATTCACTATATA  
 ATTCAGTGTGAGATGGTGGTGGTAAGTCCTATGTAGGAAGAAAAGCAGGGTAAGGAGGCT  
 TGGAGTAACTGGAGTGAGTCATAGATGGACTTGTGAGGAAAGGGTTCTGAAGAGGTGGT  
 ATTTGGGCAGAGATCTAAATAAAATGAAGCAACAAGCCATGAGAATATCCGGGGGAAAAAT  
 GTTCTGGGCAGAACATCAAGCATAGAACTTGTGGTATGATATTTATTCTAGCACACATT  
 AATTTTAAAAATGTATAAAAGACATCCATTTAATCATATTAAGATTTCCATGATTCATT  
 TAGACTTAGTCAGAAACCAAATTTATATTTTCTTTTTAAATAATTTTATCTCAACTCTTA  
 TTTTACCCAATAGGGGCCAGAGTTACTCAGCAAATACATTGGAGCAAGTGAACAAGCTGT  
TCGGGATATTTTTATTAGGTTGGTAGCCTATGAATGTTTTTAAAGTAACTGACTCTGTTA  
 TTATTTATCAATCAGTGCTTTTTTTGGTCTTGTTTTTTGAAGAACTGATATTTGAAACCT  
 GTGGTTTATGTGAATTATTAATAAGCTAGAGGACGTGGATTCTCTATTTTCATCAAATAAT  
 ACAAAACATTTTATAGATATTAAATTTTGGAAATTATTTGGTTTTGTTTTACAATAGAAAATA  
 CTCCTCAAAGTGAATCGAAGTGGTTATTCAAAGAAATCTCAGAGTAGATTCTTATATGA  
 AGCAAATAATTGCCCCCTAATTTATCTCTAAATTTTGTAAAGTTCTAAATCTTTTTTCCCC  
 CAGTTTCTAATTTATCTCTTATAAGTCAAGAGTCCATCTGGCCAATTTAATTTCACTGAG  
 TGTAATATTTTGCATATATTAAAAAACTGTATATGAATACAGAAGATGGTATTTAAGGA  
 TGAAAATAATTATTCAAATGTGATAGCATTATGGGGAGTTTTTAAATAAAAGTTACTGTT  
 TTATTTCTTCCAAAAATTTTATTATAAAGTATACAGTTAAGAGAATATACATAAAATACAT  
 ATGCAGCTTAAGGAAGAATAATAAAATGAATACTTCACTGATTCACCACCGAGTTTACCA  
 GGAAAAAGCATAAAACAAAATAAACCTCTTCCACGTAATTCCTGGGTTAAAGAGAAGTTAT  
 AGTGAAAAATATTTGGGAGCAAAACGATAATGAAAATACTATCCATTAATTTGTTAGATG  
 TTGCAAAACTGATTTCAAGGAAAATTTATAGTGTTAAATGTTTAGAAAAAGAAAAAGGTT  
 AGAAGTTAACCACTTATGTATCTATCTCATGAAATTAGGAAAATTATAGATATAAACTAA  
 AAAATATGTTAAAAGGGAAATAATAAGATAAGAATGAAGTTAATGAACACAAAACAG  
 AGAAGCTCACAAAGCCAAGATTTATTTTTGAACACCGAGTACAATTGACAAATCTCTAA  
 CAAGTTTGATTAAGAAAAAGAAAGCATGAATAAACAATTTTAGGGATAAAAAGGGAAAC  
 ATCGCTAAAGATATCCCAGAAATGTAAAGATAATAAGGGAATATTATGAAAATATTCAT  
 GCCAATACATTTGAAAACCTTAGGTGACATAGACAAAACAAAATTGACCAAAATTGAGCA  
 AAAAAGAAACAAAATCTGAGTAGTCCTGTAACCTAGTAAAAATTGAGTTAGAAAAGTTAA  
 AGAAGTCTTTACACAAATCAAACATCAGACTCAGTTTTCTAGGAGAGTTTTGCCAAACAT  
 TCAAGTAGCAGATAATTCTGGTCTATTTTTGGCCCCAGAAGATATATTTTACTTGCCATG  
 CATTAAATGAGATAGCTGTTGATTTTTTTCAATCACCGTGACAGGTGTTTTATATTAGGT  
 GTTATTCGCCAGACATCTAGTCCACCTGTTGCCAGATATGGAATTAATATTCATTATTT  
 TGAATTAATAATTTGTTAATAAATTAATAAAACAAAGTCAAAGTTCAAATTATTAATAAAG  
 TAAAAGAAATAAAATATATTTTATAGAGAGCCCTTACAAAACAGTACCAACATAATGAGC  
 TTTCCAAATTTTGAATGGGCAAAATAAATGAATAGGCATTTCAAAAAGGAAGGGGTG  
 GCCAATAAGTATATTAATAATAAAATGGTTACTTGTAATAGGAATCAAAGTGTGTTGA  
 CTTATTGACTAAGAGTCAGTTTTTGTGTTGATCCCTGTAGTCTATCCAGAAGGCATGGG  
 TCTTAATAAACACCTTGACCTCAACAGTTTACTGAATACAAGGTAATTCATATGCCTT  
 GCCTTCTTTAAGGGTTTGTGTAAGATTAAAAATAAATACATAAAATATATATAAATACAT  
 TTATATGTATTTATATGTAATTACATACAACCTGCCTTCTTTAAGGGTTTGTGTAATAA  
 TTAAAGAAGTATATAAATATATATAAATACATAAAATAAATACATTATATATGATAT  
 GAAATCACTTTGCCAACTATGAAGCCTGATTCAAATATGAAATGTTGTTTGTGTTTTCCCA  
GAGCACAGGCTGCAAAGCCCTGCATTCTTTCTTTGATGAATTTGAATCCATTGCTCCTC

FIGURE 14.11

30/64

GGCGGGGTCATGATAATACAGGAGTTACAGACCGAGTAGTTAACCAGTTGCTGACTCAGT  
 TGGATGGAGTAGAAGGCTTACAGGGTAATAATTATAAATACAGAAATAGAAATGTTATAAC  
 AAAATGTCATCATGTCATCAGATTTTGGTAAAAAATGTTCTTTTTCTCTAGGTGTTT  
 ATGTATTGGCTGCTACTAGTCGCCCTGACTTGATTGACCCTGCCCTGCTTAGGCCTGGTC  
 GACTAGATAAATGTGTATACTGTCTCCTCCTGATCAGGTGACAATTCATATTTAGAGT  
 CCAAAACCCACAAATGTACACTCTTTCCTTGTGAGCTTTACTTCTGCCAGGTAATGGC  
 AATTGTCCTTAGAAGACCAGCTTTCTTAGGGAAAAGCTTTAGCCACTGTTTGCTCAAAGC  
 ATAAAAAGATTCTGAATTAGATGCAAAGCCTTTTTTTGGCCCAGTGCAAGTCTGAAAAC  
 TTGTAATCCTTCTGTGTTGGCTGATTGGGGAAAAAATGCAAGAAACCTAATGTATTA  
 TATTTTCACATTATCTTCTGTTCAAAGATTACATACTTCCATTATCCTGTCAAAAAA  
 ACTCTGATACAGAATCAAGCATGTGAATCGTAAGCATGTAAGCAGGTTTCATAGAGATAA  
 TTTTTCAACTCTTCTTGTCTGTGTTTCCAACTCTTATTCTCCAATTAGAAGCAAA  
 CAAATAAATGAATGAAAGAACAGATAGACAAATGAATAGTCAAAGGTATAAAGTATCTGT  
 ATATATGTTACATGTAGCTATTATTTAAATTATTTAGATTTTCTTTTGAAATACCTTCT  
 TGGCACACTTGCCCTAAATCTAGAAAATAAGCACTGTGTGAATAAGAAATTATTTACTG  
 AATATTTTGTAGGTTTTTGGGTTTTTGTTTTTCAGACAAGGCTCTCACTTGTCAACCAGG  
 CTGGAGTACACTGGTACGATCACAACCTCACTGCAGCCTCTATGGCCCAGGCTCAAGCAAT  
 CTCCTCCACCTCAGCTTCCGAGTAGCTGGGACCACAGGCACAGCTACCATGCCAGATA  
 ATTTTATTATTAATTTTGTATAGAGATGGGCTCTCCTGTGTTGCCAGGCTTTCTTGA  
 ACTCCAGGGCTCAAGTGATCCTCCACCTCAACCTCCCAAAGTGTGGGATTACAGGCGT  
 GAGCCACCATGCCAGCCTTAAGAGTGTGTTGATTTTCATTCACTTTCTATATATATTAT  
 TTCTGTTGGGAAAAAATCCAAGGAAGATAAATAGTAGGCTGTTGGTACATTTCTCAAC  
 TTACTTATAAAGCTTTTTTAGATATATAAGGTTAATTTATGAAGAAAATCATAAGATACAC  
 AATTTAAGATAATATTTTAAATTTATTTTATTTGTTAAATAAATTTTTCTCCTTTCA  
 GGTGTCACGTCTTGAAATTTTAAATGTCCTCAGTGACTCTCTACCTCTGGCAGATGATGT  
 TGACCTTCAGCATGTAGCATCAGTAAGTGAAGTCTCCTTTACTGGAGCTGATCTGAAAGCTTT  
 ACTTTACAATGCCCAATTGGAGGCTTACATGGAATGCTGCTCTCGAGTGGACTCCAGGC  
 AAGTTATATAGGAAGTGTATGACATTTTATGAGTGATAAAGAAGTACAATGTCAA  
 ATTTCCACCTTAAAAATGCTATTTTTTAAACAACCTTGGTAAACTGTATAGAAACATA  
 AATTACCTTTAGTTGAATGTTCCATAGTTGGAATATGGGTTTTGCAGAGAATTTATAAT  
 TATGAAGTTGATGTCTGTTTCTTAACTTACCTTAATATTGGCAAAAACATGTTGGTG  
 TTTGCAAGGATATTTTAAATTGGGATACCATGAATTAATACTACAAACAAAAATAAT  
 TAGAGTTTTTTGTTGTTGTTGTTTAACTTTTAAAAAATAATCAGTTAAAGTTGTGTT  
 TTGAAGCTCATTGTTCCAATCTGGCCAATAGGAGCCCCTTTGTATGGCTCCTGTATC  
 TTTATGACATGTCCTCATCATTCTGAATCACTTCTCACTTCCAGATACAGTAAGTTAT  
 TCTTGGCCAGGTGCAGTGGTTACGCCTGTAATCCCAGCACTTGGCAGGCCAAGGCAGG  
 AGGATCATTGGGCTAGTTTGGAGCCAAATCATGGTTGCACAACTGTACCCACTATGG  
 ACAACAGAGTGGGATCTGTCTCTGTGAAAAATTTAAAAATTAGCTGGGCATGGTGGCAC  
 ATACCTGTAGTCTTAGCTTCTTGGGAGAGGCTGTGGCAGGAGGATCGCTTGAGTAAATCC  
 AGGATGCAGTGAGCCATGCTTGTGCCACTGCACTCCAGCATGGATGACAGAATGAGACCC  
 TGCCCCCAAAAAAGAAAAATATTCTTGGTTTATCTTGTACTTCTGTATCCCAGCCCTAG  
 CATCAGCCTTTTCTCTAAAGACAGTATTATGATTTTAAATTTTACAGTAGATATTTGAAC  
 TGTTACATTATAGACTTTACCATATATTTCTAGGAAGGATTATTCTATTACTCTCTTT  
 ACCACATTTGTTGGAAATGTCTACAGAACCTACAGTTTCTAAATCAGAAACTCCCTAGGT  
 TTTTGCTATTTTGGCAAGCCATTGAAGTTCTTCCCTCTCCCTTTACTACCAGAAAGGTGT  
 GTATTTGTAGAGCTCTCTATAATGAGAAAGCACTCTATAACATGGTTGATTCACTATTT  
 GGAGTAGAAAAGTATGAATGGAAAGTCAGAGACATAAAAAATAAGCCAGAGGTCTGAGT  
 CTTAGCTTCATTACAGACTTTCTTGGGGGATGGTTGGTAAATATCTACACATTCTATCT  
 TGTCTTTATAATTTTAAAGTAAATTTTACCATGTGCCTCAAAACCGTTAGAGAATTA  
 ATGAGCTCTTTGAAAAATGCTTCTAAGTTTCTTGTATTGCTCTAATAGAATGCTATCTAT  
 GTTATTATTTATTTCTGAGACTAAAATTGTTACATCTTAAACTGGTTGCTCTTTGTG  
 TATTTTAGGATGGAAGTTCCAGCTCTGATAGTGACCTAAGTCTGTCTTCAATGGTCTTTT  
 TTAACCATAGCAGTGGCTCTGACGATTGAGCTGGAGATGGAGAATGTGGCTTAGATCAGT  
 CCCTGTTTTCTTAGAGATGTCCGAGATCCTTCCAGATGAATCAAAATCAATATGTACC

FIGURE 14.12



31/64

GGCTCTACTTTGGAAGCTCTTATGAATCAGAACTTGGAAATGGAACCTCTTCTGATTGG  
 TATCTTGTGCAGTCATCATTATACAGTTCTGAAATATAAAGCTATATGTTGGTGTAAAGT  
 TGCAGTGATTTCTCTCCTAACAGCCCCACATATTCTTCTGGTTGGTTGGTTCTTCAGT  
 AAAATAGTCTTGTCTTCTGCTTACACTAATTGGTAATTGTCATTCTTGTAAAGATTTTC  
 AAGACAGGGCTGGGAGCAAGGAACCAAGTAGCGCGTGGTTGTGATTACCTTTGGTTTCT  
 TTGAGGTTTCTCTTACCTAGTGGCTTTAAACATCTTTAGGAGCAGTTCCATTTTATAGT  
 AAACCTAAATCTGTTATCATGAACAGTTGAGGATAATGAATAATTTGATACAATAATGT  
 AAGAAATTCCTGAAAACAAAGTGTATCTGTGATACTTTTGCTGCATAGTAAGCACAATG  
 AAGTGTAAGTATAATGTTTCAACAGGAAAGTGTGTTTATTAAATGTGGGCAGTATCACTG  
 TTCTACTAGCATTCAACATCTCTCTAAAAATTAATAGTGGTTCACTGTAAATTTTATTGG  
 TACATGTAACATCTGTACATGTGTTTGGTTATCTATATGTTTCTGGTTTGTGACATT  
 TGCTTTATTAAATTTAGGCTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTCACTCTATCATC  
 CAGACTAGAGTGACGTGGCACAATTATGGCTCACTGCAGCCTTGACCTCCTGGGCTTAGG  
 TGATTCTTCCACCTCAGCCTCCTGAGTAGCTGGGACTACAGGCACATGCCACCATGCCCA  
 GCTAATTTTGTATGTTTGTAGAGACGAGGTTTACCATATGCCCAGGCTGGTCTCAA  
 ACTCCTGGGCTCAAGCTATCTGCGTGCTTGACCTCCCAAGTGCTAGGATTACAGGTGT  
 GAGCCACTATGCCTAGCCTAACTCAGACTTTAAAAATATAAAGCAATTCATTTTATTTC  
 CCAAGAACAGTAAGGTGGTGGTTAATTTTAGTCTTTAATCTGTTTTAATTTATCTTA  
 TTTAGAAATGTCCCGAAACTTAGTATAACTTTACTTTCTGAAAATGAAGAAACCTGTCC  
 TTGGGCATTAGTGTGTTGGATTAAAGCAACAAAGTTAAAAAACCTACCCTGTGTTATGG  
 CAATTTTCACTTGATGGTGGTTCTATAACACAGGTATCAGTGAACCTTTATAAAGATGA  
 ACAACTTTTCACTTGATGGTGGTTCTATAACACAGGTATCAGTGAACCTTTATAAAGATGA  
 TTTATTGCTTAAATGTTAATTTTATATTTGGTAAACAGATAGTTTTTCTCTCCCCC  
 TCTTCTTCCATCTTTCACTACTACAATTTACCATGCAGAGCTCACAAATGTCTCTGCA  
CCAAGTCCATGACTCAGGATTTGCCTGGAGTTCTGGGAAAGACCAGTTGTTTTACAG  
CCTCCAGTGTGTTAAGCAGCTTCAACAGGGGTTGCCAAGAACTTACACAAGAACAAAGA  
GATCAACTGAGGGCAGATATCAGTATTATCAAAGGCAGATACCGGAGCCAAAGTGGAGTA  
 TGCTTTTTTCCCCCTCATTATAATTGTTAAACTTCTTAAAAATGTTTCACCTTTTGA  
 TATATATTTCTTTGACTTATAAACGAGCTATATTTATAAACAGGGACCAGAACACATTA  
 ACTCAGTCATGGTTATGTGCTTCTTGGCTTTCAATGTTTCATTATCTTATAAGGAAGAGA  
 ACGTATGGTCTCTTGAACAACTGACAATAAGAAGTAACAACCTGGACTACCACATTTTT  
 TTTACATCTTAATTTAACTCTTCTGCAATTTCTTTTTTACTTAAGGAGGACGAATCCA  
TGAACCAAGGCAGGACCAATCAAAACAGAGCTGGCTATTAGTCAGTCACATTTAATGACTG  
CACTTGGTCAACAAGACCATCCATTAGTGAAGATGACTGGAAGAATTTGCTGAGCTGT  
 AAGTAACAGATTCTGTTTTGGAAGTACAGCTACTATTACAAGTGACATAGTATTACACTT  
 AAACCTTTAAAGTTCGTGTTTAAATAAAAAATATTTGAATATTTAAAGCTAATTCAAA  
 AAATATGTGTCGTAGCTATGCATTAAAAAACCCCAAAATGTGAGAAGTACAGAAGTCAAA  
 ATTGAGTTTTCAATTAACAGTTCAATTTGATTATATTTGAATTATTATAATGGACTCATT  
 TAATTTTAGTAATTTGGGCTGGGTGCTGTGGCTCATGCTGTAATCCCAGCTCTTTGGG  
 AGGCCAAGGCAGGTGGATCACCTGAGGTGAGGAGTTGAGGCAAGCCTAACCAACACGGG  
 GAAACCCCATCTCTACTAAAAATACAAAAATTAGCCAGGTGTGGTGGCATGTGCTGTAG  
 TCCCAGCTACTTGGGAGGCTGAGACAGGAGAATTGCTTGAACCCAGGAGGTGGAGGTGTC  
 AGTGAGCCGAGATTGCACCACTGCATCCATCCAGCTGGGCCACAGAGCGAGACTGTGT  
 CTCAAAAAAAAAAAAAAAAAAATTAGTAACTTCGAAGAAATAAGAAGGAAATTTAAAGT  
 TGAAAGTGATTCTAATGTATAGTTTATAAAATTTTGTATAAAAATACCTGTTTGCCTT  
 CAAAATAATTTATATTAATTTTATTGACCTCAAGAACATTTAAATACATTAGATTTA  
 TTCAATTTGTGGACCACATTTGTTATACATTGGATTAAAGGATCCTTGCAATTGAGTTTA  
 TGGCCACCTATGCATCTGAGACCCATGGACTGGGAACCATCTAGGTCAATGATTCACTG  
 TGATTCAATTTAAGAGATGTTTATCTGGTCTTTAGAAGCTGTACCTTTTGTATCTA  
 ATTTTGCACTTTTGAAGTATGTATGTATGTGTACATACGTTAGTGTATGTATTTATT  
 AAAGAAGAAATCAGAAAACAGAGGTAAGGAAAAATAAGGAACAAATTTCTGTTAAGCCCA  
 CCACCTCCCAAAGCATATTTGTTTATATGCTTATATATGTTTTCCTATTATGGTAAGAAC  
 AGTCTGTACATATTGCTATATAGCAGTCCCCCTTTATCCACATACATCCTGAAAATGTT  
 TTACATTTTAAATGTTAACTACTTTATTGTTTTTAAATGTCAATTTATAGTGTAGCTATG

FIGURE 14.13

## 32/64

CCACAATATCCAATTTTTAGACATTTAAATTGCTCCCAGGCAATGTGGTAATGAACATTCT  
 TTGCAGCTGAATATATGCACATATCTAATTGTTTCACTAGGATAGAGGTGGAATTGTATA  
 ACAGGGAGCTCACATTTTTAAGGCTTTTGAATGTATTGCCAAATTGCCTGCCAGATAT  
 ACTGCACCATCACTAACATTGTGTGTGTCAGTATTTTCTAAACTTGGCCCTTTTGATTT  
 TAGAAAAATGATATCAATAATTTACATTTCTTGATTAAAGTGTAGAAGTTATAATTTT  
 CATATTATTCATTGTCAATTTGTATTTTATCTTTTCTAACTTGTCTCTTCATCCCCTTTGC  
 TCCGTTTTCTATTGGAGTGCAACTTTATTTGTAAGAATTCTTTTAAATTTCTGTGACTGG  
 AATTTTTTTTTCTAGTTTGTATTTCCCGTTCAATTTCTAAAAATAAATTGTGTTTGCCA  
 ACAATCCATTATCTTTTGTGTAATGGTAGTATTTATACATATTAATTTATCTCTTTC  
 TTTTTTCAGATATGAAAGCTTTCAAAATCCAAGAGGAGAAAAAATCAAAGTGGAACAAT  
GTTTTGCACCTGGAGCAAGTAACCTTAGCATAAAAATATACTTCTTTTGTATTGGTTCT  
GTTAAGTTTTTTGATGGCTTTTCCATATGTTGTAAACAGGAAAAAATGGTGTCTATGAAT  
TTCTTCTTAATTTAACAAATTTGGTTAATTTATAAAATCACAGATTGGTAAATGCTATAA  
TTATGTAATGATCAGGATTGAGATTAATACTGTAGTATAAATTGGGACATTATAACAGAT  
TCCATATTTTATTTCTAAATCTAAATTCAGTCTTTAATGAAATAATATTAGCCAAATG  
GTGGAACATAATTTCTTTTGAGGAAAAGATAATAAAGAAATGTAATTAATTTAAATTT  
TCTTGAATTTCCAGTTGTATATTCATCACCTTTGTAGCATTGTGACAAATTTATGCTTA  
GCAGCTTCTTCTCACTGTTTTGAAATAAAATATCCTATTACCTACTGATACAATTATCTGTT  
 CTTTGTATATCAAAAAATGTGAAATTTACACATAATTCAAATACATTTAATTATCCGCTC  
 AACCAGAAATGAAATCACATCCCTCTACTATACTACATCCAGCTCCAAGCCCAAGATATT  
 TAAATGACATCCATTCTCTCTAGTTCCAGTTATGATTTTATCTTGATATTCTCTCATA  
 TATGAATAAATATAAAGTTAGCCACCATCAATACAACTGCGTATCTAATATCTTAAC  
 TATATAGTAATGGGGTAAGGGAACAGCAAAAGGAGAACATTAATTAATAATACAAAGTA  
 AGCCTGGGCAACATAGTGAGACCCCATCTCTAAAAAATAAATAGCCATGCATGATGGT  
 ATGCCCTCTAGTCCCAGCTACTTGGGAGGCTGAGGTAGGAGGATCACTTGTCTCCAGGAGG  
 TTCAAGGTTCTAAACCAGCAAAGCTCAGAATCCCAGGGGATAGAAACAAAGACTTAGTGG  
 ATCACTAGTATTAACTGAGACACGTCACCTGCATTGCACTTTGTCTCAGTTCTTTG  
 ATGAAATCACTGAGCTGACATACCTGCCCTCTTTTACCATAAAGTGAGTTTCATGATCA  
 GAAGCAATGTCTATGGGATAGCCTAACAAACAATGTAAAAACCATTTAGTAAGTTCTATGA  
 AGGGTGGTGGTGGTAAAAATTTGGGAGACATACAAACAAATACAATTTCAAGGTGTGTC  
 CCCTCCAGGAAGGACAAATTTGCTGCTGCTCTGTGATAGAAGAGGATCAGATGTAATCAA  
 CTTGCCCTCAGACTTGGGCTGTTCTCTCTGGGTGTGGACTTGCTTGGTTGGTCACTGCT  
 GCTGACAAGTAGGCTGTCAATATAGCTGGGTGTCTGTGCTGCTGTTGGTGAGGGGGAAGT  
 CCACATTGTGGAGGCCACATCCCTGCACTCTTGCCCAATTTGACCATGAATCTTAAGCAC  
 TGGGGTGGCTGGAAAAGACAGCCGATTGACATCCATACAGAGGTCATCTTGACCACTTGA  
 TTAGTATAAGCACTGAAGGCTTTTAACTGAGCATTACATAGGACACAAATATTCTGATT  
 CTTTGGGCCCCATTCCAAGAACTCTGGGCATCTTTTCTCCAGACCTCATACCCAGTTGT  
 GTTCTTTCCAAATTTCTGGTCATCTGGTTATGTTATTAGCCACTATCTGTGAATCAGCAT  
 AGATTTTTTATATCAGACATCTCTACCTCCTGACAGAATGGAGGAGATATGTTACTTAACA  
 ATTCTGTTCCCTTGGAGATTTCCTGTCTCACTGTTTGTAAAGGGCTACTCCCTCAATGT  
 AGCAGTAATGCTTTCACTCTGATGGGAAGTCACAGTGGAATTCTGGGTCTCCAAGAATTA  
 GTGTTAGTGCATACACAGTGTCTGATAATCCCAGAGTGTCTGGTGGCCCTTGGATCCTGT  
 GAAGAAGGCTTGGAGAAAAGAAGATTCTAGGCAAGAACTTGTGATGTGATGACAGGGCCT  
 TTTCTCTGGCTCTTCACTCTTAGTCTGACCTAGGTGTGAGAATTAGGTCAGGGGCCATGA  
 CTATATTGTGGTGACTCAAACCAGGCCTTTGTTTACTAACTGGGAGATTTTACATTGTA  
 AGAATCAAGTAGGATCTTTGCCCATGTATTTGGTCTTAAGAACACAAATGATATGGCTC  
 CAATGACTGGAGGAACACCAGGGTCCTTGGTCTCACGCTGATTTAGATAAAACGACTGTC  
 AGGCCTCTGAGCCCAAGCTAAGCCATCTTCCCTGTGACCTGCACGTATACATCCAGATG  
 GCCTGAAGTAACCAAGAATCACAAAAGCAGTGAAAAATGGCCTGTTCTGCTTAACTGA  
 TGACATTCACCAATTGTGATTTGTTCTGCCCCATCTTAACTGAGCGATTAACTTGTGA  
 AATTCCTTCTCTGGCTCAAAACCTCCCCCACTGAGCACCTTGTGACCCCCGCCCCCTGCC  
 CCTAAGAGAAAAACCCCTTTGATTATAATTTTCCACTACCCACCCAAATCCTATAAAATG  
 GCCCCACCCCTATCTCCCTTTCGCTGACTCCTTTTTTGGGACTCAGCCCGCCTGCACCCAGG  
 TGAAATAAACAGCCTTGTGCTCACAAAGCCTGTTGGTGGACTCTCTTACACGGAC

FIGURE 14.14

33/64

AAGCTTTAGTAGAGATCTCAAAAATGGTTGGATGGTAGCAAATTACTAAGAACTCTCAA  
 GTTCTAAAGCCTTAGTTTCAGCTTGCTAGAAAACCTATGTTGAGTATTATGGCTAGTTC  
 CATAGTTGAGTTGGGAAATGTCTTTGAGGAGACACTTTTTCACTTTGTATTTCATCTGTAC  
 ATTTTCTGTTACTTGCATTCTGTCTAGCTCAGGCTATTAGAGCAGGTACATTTTTATAAC  
 TGGAAATGTTTATGTGTAGTGAAGCTCTGAGAGGACTTTGCATTAGATCTCAGCAGCATAA  
 TCAGAAGGTTGTCTTTGTCTCAGCAATTTTAAGCTAATAGTAGCAGAAATTGCAGTGG  
 AAATAGACTGCTTTGCCACAACATTAGAAAATCATTTATCTTTTTATTGCAGTTCTTGT  
 CACCAAACAATACATTTTAGTACTTCTCAAATTGCAGAACTCTCATAGGGCTGGGAAAAT  
 GCCTGTAGACACATACATATGAATGTGCTAATGTTTTTGTATTTTCATAGCCCATC  
 AAAGCTCCTGAGTCAGTTTCCACTATAATCACTGCAGAAATCAATCTTCTACAAGGTAAGC  
 TTTTGTAGAGTTACTGAAGGAAGAGTTGGGCCTAGTGGGTAATGTGCCACTAAAATGTTG  
 GATTAGTCTAAAGGCTCTGTCTACTCTTTATTTGTATAAGGTGTGATTATACTTTTTGTT  
 CCCTTCTTAGCTGTTTTCCCCATAAGTGGCTGTTATTAACATCTCATCTAGAGCTGA  
 AGTGGGAGGAGAAAGTGCCTACTGACACATGATGTGAGGATCTTAAGTATTTTTTTTAG  
 TGTAGATTGTAGGAATTATTTCTAAAATGCTGATTGTATAGTGTGGAGCCATGGAAGACT  
 GAGCCGTTAGTACGATGGCATTGAAGAATGAGAAGGACAGAGACAGGATTGGACTAGTA  
 GAGGTTGTCGACTGTGGTGTCAAATGGGTAGAGTAGGCCAGAGATTCTAAAATGCCTTT  
 AAGTGGAGTTGAGCTGAGTAAGGGCAGTAGTGAGGATTAACACCTACTAGAAATTCATAG  
 TGAGAGGAATTCAGATGTTTTGATAAAAGAAATGAGGAGGTCAGGTTTCCAGGGCCAA  
 AGTCCATGAACATCTGATACCTCAGTGAGAGAAGTGACAGATTGTTGTGTTTAAACCAGA  
 AGTCTTAGGAAAGGAATTAGAACATAGACCCCCAAGGCTCGGCAGGCCTGGCACGGCACA  
 GGCAGCAACCATTGAAGGCTATTTGGTGTTCGGGATCTGAACGTCTCATTTAGGGGACAG  
 TGGTGTGAGTTAGTACTTTTACTTTGACCCAGGTGGACTGAGAACTCAAGTGATGATGC  
 CCTTAAGTATACTTTTTTTTAAAGCCACAATCTATATAGTCGAAGTCTGTTCTCCCAAC  
 AGGGGTACACTGGCATTCTCTCAGCAGGGCTGGGAAAAACCAACAACAAAAAAGTCTGTA  
 CACAGGCAAACATCTCTTATTTTTCCAACATTTAATACATTGTTAATAAAATATCTAA  
 AGTTTAGCAAACAGTTGCTGTGTATCAGTGGCTGAGCATTTCATGCTTTTATTTTCATTC  
 AGTTCACCTATGAGGTGGATACTACTATCCCCATTTCTAGATGAGAACATTGAGGCAC  
 AGCGAGGTTAATTAACCTGTCCAAGATCACATAGCCAACAAGTCATGGAGTGAGGCAGTC  
 TCATGCCAGAGCTTAAGCCTAGAGCATAGTTCTGGCTCTACAGCTTTAGCAAGTGACTG  
 GCTATGTGACGAGGACCAACCTCTCTAATGTCTCATCTGTAAAATAGGAATTGTAAATAG  
 TTACTACCTCAGTGGGTCAAATGAAATCATATGTGTTAAGCACTTAGCAGAGTAAGCACT  
 CAATGAATAGTAGGAGTTATCACATCTTCGTATTTGTGCATTACCTTCACAGTTTACAGA  
 TTAAGGCCAGAAGCAACTGTTGAGCTACGGGTTAGTGTACTAACAGTTTCCATGTGTG  
 TCTCCATGGAAGGGTGTGTGGGACCTGTTATTGTGACTGTCTGTACTTTTCGTATTGTTGT  
 CTGCCACCCATGTTTATTAATGATAAGGACAATAATGCAACAAAGTAGTCAAGTAATGT  
 TGCAAATGCCAGTATTGTAGTGGCTATCACAGCAGTGCCACTGGCAGGCAGCACCATGG  
 TGGCAAGTTCAAGAGGTCAGTCCAGCCACTGAGCTAGAGCCAGATCAGGCATGCAAGA  
 GGAGCCTGAGTGGGAGCCACTGGGGATCACGGCCAAGAGTGTGACCACCCAAGACCCAGA  
 ATGGCTGAGTGGCTCCCTGGAGCATGGCAGTGCCAGAACAACTCCATGAACCTCAGATCT  
 GGTGATGCCTAACTAGTGTGTTCTCGTGTGGACCCCTTTCTCTACCAGAAACCTTGA  
 ATCCTCTCAGCAAATGAGGAGACTACTCAGATCAGTGACTTAGTCTGTTTGGTGTTATA  
 TATGTGTACACAACACAGCACATATTAATAAATACCTACTATGTGCCAGGCACTGCCTAC  
 CACTGGAATCTTTCATAAGACATTGTTTTACTTTGCATTCTGCCTTTACACTATGAA  
 AGTAGATGTTTTGGATTATATTTCATTAGCAGATACATTTGAATATGCTGTGTTATGCATA  
 GTAAGCCTATGATAAGCAAGTATTCTCATTTAGAATTTGGGAATATTGATTATACATGTG  
 GACAAACAAACCATAAATGCAAACTATTTATATGATAAATAACTTTGGACTGATGGCTGG  
 GAGGAAGGACCAGCTATTGATGGGTAGGAACTAGCAAGTAGCGGACTGTGGCCTGCATAG  
 ACCAGACCCATCCGTAGTGATCCAGATGAAACAGCCACCCCTCAGACACTGGATAAAGGG  
 TCCACCAGGAAAAAATCCTGGCCTATCAGGTGCTATGTTACAGTTCAGTTACTGGAAGT  
 ATTTCTCAAAAGTGTTTTTATGGTTGAGGTACACATTCTACAGCTTTACCTGCTGCCA

FIGURE 15.1

34/64

AGTCCCTGTTTCAAGGGAAGCAGCAATGAATTACACTGTTCCCGTAGTCAAGGACAGTAT  
ATCTTACCAAGAACTATACCCACTTAAGGAGGTGCTGGATGTCATAAAGATTGGATCAA  
CCATTATGGGTGTTTCAGAGGAGAGATTATTTCCAGCTCAAGACCCAGGGAAGAGGACATA  
GGATGGATACCAGAGTCATAGGGAGGATTTAACACAGGACATGTACACATTAGTTAGTTG  
GGTATAAAGTGAACAGAAATGAATGAGACACAAAGCCTTGAATGCCAGAAATACTAGTA  
GTCCTGTTGTGGAAGGATATAAACTCAACTGGGAGTGGAAGAGAAAGGCAGCAGTGAGT  
CTAGGAGATGTACAGTAGGTTGAGGTAAACATATCCTGAAGACTATAATCCAAAGATTAT  
TTTTGGTTTGAATTTGTTTGGTTTGAATTTCATGGTATCTATTTCTTTGAGTGGATGGT  
TGGGGAGGGTGGCATGTAGAATGCATTCTTACCAAATCAGCATGATTTTCAAGACAGTAC  
AGAGAAAAGATAAAGCTGATGTAGGAGCTTTGGCTGCAGTCTCTATGGCTTTCAGCA  
AGCCGTTTAACTTACTACTGCTTCATGACTGTGGCTAACAAAGTAGGGATAGTACGGAG  
CACAGAGGATTTTTAGGGCGGTGAACTATTAATACTCTCTTTGTATGATACTATAATGG  
TGGGTACATGTCATTATACATTTGCCCAACCCACAGAATACACAGCACCAAGAGTGAAC  
CCTAATGTGAACCTCTGGTCTTTGATGATGCTATGTGAGTGTACGTTTCATCCGTGTAACAA  
GTGTACCACTCTAGTGGTGGGAGGGGTATTGATAATAGGGGAGGATGTGCATGTGTGGG  
GGCAGGAAGTATATGGGAAATCTCTCTACTTCTGCTCAATTTGCTGTAAACCTAAAACC  
TCTGTAAAAAATAAAGCTCTATTTTAAAAAGTGGGGATGGTATTACGGCAATATAAAAT  
CAAAATACTTTATGAACAAATCTTTCTCCAGATGTAACTGTCTATATGCACCCCTCGT  
ATGTGTATGTATAATTTTCATTCAAACGTGAAACAACCTTTAGAATTGGCACCAACATAT  
AAACACTGATACATTAGACTATCTCGAACACCTTTTACTGACCACTTGTAAACCTTGCTT  
ACCTATTAAAGGTTCAATTCATAGCTGTGATGTTCTATTTTCAATGTGGGATTATC  
TTCTGTTTCCCCCAGGGAGTATATTACCAAAATTGGTGATGTTGTTTCTGTGATTGATGAA  
CAAGATGGAAGCCCTACTATGCTCAAATCAGAGGTTTATCCAGGACCAGTATTGCGAG  
AAGAGTGCAGCACTGACGTGGCTCATTCTACCCCTCTCTAGCCCCAGAGACCAATTGAT  
CCCGCTCTATATCATAGGTAAGTTTGACAAATGGCACAGGTTTTTTTTTAACCTTAGTT  
AACTCTCCAATATTATGTAAAGAGTGTGTTAGTCAGCTTGGGCTGTGAGGACAAAATAT  
CACAGACTGAGTGGCTTAAACAACAGAAAGTCACTTCTCACAGTTGTGGAGGCTGAAGT  
CCAACATCAAGGTGCTGGCAACACGGATTCTGGGGAGGCTTTTCTTCTGCGCATATAGA  
TGGTCACCTTCTGTGTGTCTCACATGGCCTTTCATGGAGTGAGAGCTCTTTGGTGTA  
TCTTCTTATAAGGACACCATTCTGTGATGAGGGCCCCACCCTTATGGTTTCATTTAA  
CCTTAATTGCCTCCCTCAAAGTCTCATCTCCAAGTACCATCACATTGGGGATTAGGGCTT  
CAACATATAAAATTTGGAGGGTGGCGGGGGGGATGCAATTCAGTCCATAACAAAAAAGC  
ATGAGTATTATTAAGTACAAAAAATTAGAGAGCTTTATAGAAAATATGAGGCATTTTAT  
GTAGCTGGAGTGTGAGTGCTATCAGTTATTTGAGTTAGAGCAATGTGCATCTACTAAGA  
AGTGGTATGGATAAGATTTTTTGGAGTGACCCAGGGTTAACTGTACTACAAGAATGTA  
TTGCTCAGGAACCTAGGTTATTTAGGTTACTTATTTATACAAACCTATTCAAAAAATAATT  
AGGAAAGAACTATCCAGTTATCCCATACTTGCAAATCTCAATATGTGTGCCTCTGCAT  
GCTACACATGTCATCTTAGGCCTTTATAGTATAAAGGCTGATAGTTGAAATGGCAGCTGC  
TGTGCTTTTGTAAATTTCAAAGCTGCCAAAACAGTTGTGAGATAGACTCACAGAATTTA  
CTGATTAATACAATTTTTAAAGTTTTTCAGATTTTTACAGTTACTTCAGACTTTTTATCTT  
TCTGCAGTGAGCATGCATCATTACTTTTGCATCCTGAGAACAAGCATAAGTGTGTTTTG  
GAGAGAACTCCAGGGACAAATAATATACCACTGTTATTCTCACCTATATGTCAAGTTTGA  
TACATTACCAAAACAACTTAGCCTTCTGCTTATAAGTATATAGAATTTTTATTACCTTA  
TCTATGGATCAGGATCTCAGCAGAGGCAGTGATGTATCAGAATCACCTTCGGGATTCCTC  
TACTGCCCTCTTTCTAATCCCCAGATTCTGATATGCATCCTTGTCTACAGCGAGGCA  
GCATGGCATGAGGTGAGAACACACAGTTCTGGAGCCAGACTGTCTAGGTTACAGCCTGCC  
ATTTACCGCCATGTGACTTTGGCAAGTTTCTTAGTCTCTCTGCCTCACTTTCCTCATA  
TGTAATGGGAATAATAATAGTGCTACCTCAGAAGGTTGATGTGAGGAATGAAGGTAT  
TGATACATGTAACTTAGAGCAGTGTTGGGTACAAAATAAATCATGATGCAAGTGTCAATC  
ACTGTTTTTGGGAGAAATGCCATATTCTTTAAGCCGTTAAAGAAGAAAAATGATTAAGAA  
TAATTTCAAAGTAATGCATGTTTCAAGGGCTAATGCCAGGTTGCTCCAGAGTGGTCTCT  
CCCAGTGCTAGAAATTTAACATCTTATGAAAATGATATATATGGTCAAAAATGTATTT

FIGURE 15.2

35/64

AACCTTTCCCTTGGCTGCCTTCCAGGGCCAGAGGAAGATCTTCCAAGGAAGATGGAATAC  
 TTGGAATTTGTTTGCATGCACCTTCTGAGTATTTCAAGTCACGGTCATCACCATTCCC  
 ACAGTTCCCACCAGACCAGAGAAGGGCTACATATGGACTCATGTTGGGCTACTCCTGCA  
 ATAACAAATTAAGGAATCAGTTGCCAACCATTGTAGTTCAAAATAAAACCTGGGTTTCC  
 AGGCCTGGTGTGGTGGCTCAGCCTGTAGCCCCAGCTATTGCACCACTGCTCTCCAAGCT  
 GGGCAATGGAGTCAGATTCTCTTTCTTAAAAAACCAAAAAAACTGGATTTCAGTTCT  
 CTAATATTCTTAGTACCACAAGATATGTCATAGGTATCTTTAAATGAAATCTTAGCTGG  
 AAAAGTGACTAAAAAGTTTTCTCTGCTACCTAGTAATAAACAAATCATTGTTTATTAC  
 TGGTCACTTAGAAAAATTAAGGGGATAGGGCCAGGCACAGTGGCTTATGCCTGTAATTGC  
 AGCAATTTTAGAGCCGAGGCAGGCGGATCACCTGAGGTCGGGAAGTGGATCGCCTGAGG  
 TCAGGAGTTTCGAGACCAGCCTGGCCAAACATGGCGAAACCCGTCGCTACTAAAAATACAA  
 AAATTAGCCAGGTGTGGTGGCATGTGCCTGTAATCCAGCTATTTGGGAGGCTGAGGCAG  
 GAGAATCGCCTAAACCCAGGAGGTGGAGGTTGTAGTGAGCCAAGATTGCACCGCTGTGCT  
 CCAGCCTGGGCAACAGAGTGAGACTCTGTCTCGGAAAAAAGAGGCTG  
 GGCACAGTGGCTCAGCCTTAAATCCAGCACTTTGGGAGGCTGAGGCAGATGGATCGCC  
 TGAGGTTGGGAGTTTCGAGACCAGCCTGGCCAGCATGGTGAAACCTGTCTCTACTAAAAA  
 TACAAAAATAGCCAGGTGTGGTGGCGCACACCTGTAGTCCAGCTACTCGGAGGCTGA  
 GGCAGGAGAATTGGTTGAACCCAGGAGGCGGAGGTTGCAGTGAGCAGAGATCGTGCCACT  
 GCCTCCAGCCTGGGTGGACAGAGCAAGACTCCGTCTCAAAGAAACAAACAAAAATTA  
 AAGGGATAGAATATAATGAAATATATTTGAACTTAAATTATATTCTATATGTGTATCTT  
 CCTAGGCAAAAGCTGTAATTTCCAGAGAGACCATTAGGAACAGGTAGTATCTATTTTCT  
 CCATTATTTATTTCTAGAACTCATAAATGGATTGTATTTTCTATAAGAACAAATAT  
 TAATTAAGGTATAGATGACTGACCAAGGGCTTAATCAAATAAAATGACTAACAGCATCTA  
 TCATAAAGCCACACAAGCCTTATGTTCTCATCTCAAAATGCTGTGACAGCTTTTGGCT  
 GCTTTAACCATAAGAAAAATGATTGGTGGATGATTTATTAGCCAGGCTTTAAAAACT  
 TTCATCTAGGCCACGTGCGGTGGCTCATGCCTGTAATCCCGCACTTTGGGAGGCTGAG  
 TGGATGGATCACTTGAGGTCAGGAGTTCAGGACCAGCCTGGCCAAACATGATGAAACCTG  
 TCTCTACTAAATATACAAAAATTAGTTGGGTGTATGGTGCATGCCTGTAATCCAGCTA  
 CTCGGGAGGCTGAGGCAGGAGAATTGCTTGAACCTCGGGAGGTGGAGATTGCAGTAAGCCG  
 AGATCGTGCCACTGCACTCCAGCCTGGGTGATAGAGCAAGACTGTCTCAAAAAAGAAAA  
 AAAGAAAAAATTTTAAATTAATCCTTCTGTAGAAACAGGCATTCAGAACCATTCCATTGA  
 TCTTAATAAAGCTGTCTTTACTGTTCTAGTCAAAATGAGACTTCGATCAAAACCATAA  
 GATTTTATACTGCAGATAGTCAGCTTCAACAAAGCCGAGAGGAAACATGTCGAGATCAG  
 GCTTCTGCTTGATAGTCTCTTGACTACCATTAAACGAATATTGGGAGGTGATGAAAGT  
 CATTGGTAGGCCATTAGCATGATATCTTTAAACATCTACCCTAAACCATCTGCTATGG  
 ACCCATAAATAGAGGCCTGTTGTATATGAAATGTCTAGAAATTCAGGTGCAGGTCTTTGC  
 CGGTTAAGTAAGGAGCAACACGTAATGGGAGAGGAGTGGGGTGTACTCACTTGCCCTC  
 CTCTTTTGTCCTGATTTAACCAGCATTTTCAACCTGGGAAAAATTTGCAGAACTAAGT  
 TGATTGTAATGATTTTGAGCTGCAGCAGCTTTAACTCTTACCCTTTTCCACATAGTTAT  
 GGTGTTTGAGTTGGAAAGAAACAACCTATAGGTAGCTACACGTACATAATTATCTCTTAT  
 TCACAAAGGGTATAGTAAAATTGATTGTAAATAACTTTCTAAGTGCCAATATTCAAAACT  
 TTTGGATTAAAAATGATTTTTCACCGTGCATTACTTTGGATGTATTTATTCATTAAAA  
 CAATTTAAATGGGGCTCTTTAACCAAAAATGGTATTTAAACCAAAACAGTATCGTACTT  
 AGAATTTGGAGTAGAGGCCGGGCACAGTGGCTCAGCCTGTAATCCAGCACTTTGGAAG  
 GCTGAGGCAGGCGGATCACCTGAGGTGAGGATTCGAGACCAGCCTGGTCAACATGAAAC  
 CCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCGTGGTGGCGTGCCTATAATCCCA  
 GCTAGTCTACTCGGAGGCTGAGGCAGGAGAATCGCTGGAACCTCAGGAGGCAGAGACTGC  
 AGTGAGCCGAGATCGGCCACTGCACTCCAGTCTGGGTGACGGCATGACTCCATCTCCAA  
 AAAAAAAGATTTTGGAGTAGATTTCATCATTAAATAAGTAACAGATTTTAGGAAA  
 ATCAAAAAATGGCTAATAAAATGAACACAATGTAACATTTTATTAATGTAGACTTTT  
 AAAATCTATAAATGATCATCTGTTTATAAATGGCAGATGGTTGTGTACCATCTTTA  
 AAATAAGATTGAATTTACCCAGTGTGATGGTTCCCATGCTTATATTTCTCTGCTGA

FIGURE 15.3

36/64

GGCCGGACCTGATATGGCCCTGGTCTGTGTTCCAGCCTTGTTTCTCATTACCACTAAA  
 ATCTTTCCCTGTATGCCCGCCCAATTTTCTGGCTCTGAGTCCTTGTTTCATACTGTTCT  
 CTCCAATTCTACCTTCAAAGGCCTTTCTTAACACCTTCGGATTCTTTCTTTGAGAACTT  
 TCCAGATTCCCATGCGCTTTTGGAAATCAATCTCTATCCTATTGTCATCACATTTAAGTTT  
 CTACTTCCATCATCCTCACTCCTATCCCTTTGGTCTGGGATGACAGGGATGCTGTGTTT  
 TATTTACTCATCTTTGTAACCTCCACATAACCTAACCCCGTTCTTGCTTATGGGAGATG  
 CTGATTGTAGGGTCTGAGTTAGATACTGTTAACTAAAATGCTTGTTGATATTTTAGTTAT  
 TAATTTCATTTAACTTTGGCTGAACTTTTAAATCTATTGTGAATAGTCAAGTAAATTT  
 TAGATTGTTACATTCTGGGTTAGTATTAGATTGTTTAAAGATTGTTTAAACAAGATGT  
 TTTAAAGATGAGTTTAAATAGTTCTCTTAACACAAATAAAGCTTAATATGAGTATTTGA  
 AGGAAATTATCCCAAACCATTCAGTTCTCTGGCTGTGAAAGGCTTTTCCAGGCCTAATAA  
 GTTTTCCACTTCAGCCGTAAGTAGGTGAAATCAAATGAACAATAGAGGGAATGTATTTA  
 TTGCTTTATACACATGCATGTGTGTTGTGTCTACATATAAACATTGCACACGCTTAGAA  
 TGAAGTTTCTGTATGCCAGAAAAGGAGAGGCATTTTGTGGATTTTGTCTGGCTGCC  
 CTGGGGATGTTTGAAGAACTGTGCTGTTTACTTCATACCAGGTGTGTGAGCCATACCTTT  
 GGTAGGAGGTATCTCTACACCCAGAAATATAAGCCAGGAGAAGGTCTGTGCCAAG  
 AGAAGGAACCCAAATGACCCACAAGAGGTGGGCCATTAAATTATTGGGTGAGTGCATAAA  
 TGCACAGTAATTTATTTAAGCACCTCTTAATGGTGACCCACAAGGAAGATTGCTCGTAGT  
 AGCGGAAAGGTTACAATAAATAAGAGAAAAAGCAGAATGTAGAACTGTATGATAGCAA  
 TTCTGCAACAAGAGCATCTTTTATAAAGATGGAAGGAGCCAGGCACAGTAGCTCAT  
 GCCTGTAATCCCAGCACTTTAAGAGGCTGAGGTGGAGGATCACTTGAGCTGCAGTGACCC  
 ATGATTGTGCCCACTCCAGCCTGGGTGATAGAAGTGAGACCTTCTCTCAAAAAAAAAA  
 AAAAAAAAAAAGACGGAATTCCTCCAGAATTTAACATGTCAACAGAGGTTTCTGTC  
 AGCTACTTTTTCAGCTTTTACTTTCGAGTATTTTCAAATTTTCTTAACAAGCAGTA  
 TTTTCAAATTTTACAATAAGCACACACACACACACGTTTGTGTCATAAGTGCCC  
 AACTGGTGGTGAACAACCGCTGGCTTTTAGTCTATACATATCTAGAATATTTTATAAATA  
 GTAGTTCTTAAACCTTGAAGGGAGTGAATGACCAGCTGAGAAAAATAAGTCAGTGATT  
 TCATTATTTTCTATATTCACATCATGATTCTAGGAAAGAACTTGGGAGTGACTTCCTTC  
 AGCTTCAGCCACTCCTGGGCCAGGCGCATGCTTAGCTCTGTGGTAAAGGTCACCAGCTTC  
 TTCTGCAGGGTGCTGTATCATCTGAATTGGAGGTTTGGCGAGGGTAAGAGACTGATGTA  
 GGTTCAAGTTTCTTCTTCTCTCCTCCACTTGAAATCTGTCTTCCCTTCCAGACTGCCTG  
 CGCTGCTGACTTAAGGCCCCAACACCAACACAGAAGCAACAGCCTTACACAGAGTGTTT  
 AGCAAGCTCCAACAATTGTGTAAGGTAAAGTTTCTTTATAGATTCCTTTTCTATATCGC  
 TCCTAGTGGTTCTGTTTCTCTGATCGAATTCTGGCTGATAACAGTTGCTGAGACTCTGAA  
 AGAGAAGGCAAGGAAGTACTGTTTCTCATTATAAACTGTTTAGAATTATTTGGCCATCTT  
 TTTGCTATGAATATGTAGTGCTTTGATACATTTTAAATCAAAAAGTAATGAAAGAGAT  
 CACATAGGGAAAGATAGATTGGATTATTTTAAAGTTTATATACTAAATTGAAAAGCAAA  
 GAATAAATGGGAGAAACAGCTCCCTCATGTGGCTGTTGGCAGGAAGCTTCCATTCCCTCT  
 CTGTGGGCCTCCACAGGTTTGTCTCACAGCAAATGGTCCGTGACAGAAAGACGCAAGGGCA  
 GTTGCACCCAAGATGGAAGCCACCATCTTTTCTATAACCTAATCTGAAAGAAGGGACATA  
 CCAGCACTTCTGCCATATGCTGTTGGGTACACAGACCACTCTGGTACAGTGTGAACAC  
 AGGACCACACAAGGGCGTGAATTCAGGGGACAGACCACTAGGGACCACTCAGAGGCA  
 CAGAGGGACACCCATCCAGCTGGTGGCCAATGTAAATTAACATAGCTTTTGAATAGC  
 AATATGTATCTATAATCTTAAAGTATTAAGTACTTCTTGATCCAGTAATTTCAATTC  
 TAAGAATCCATGCTAAGAGGATTTAAATGTGGACCAAAAAATGGGTATAAAAAGAGTT  
 GTTAACAGTATTTAAAGTTGTGAAAAACCAGAAACAATCTAAAGGTCCAACAATAGGAAA  
 ATGAATTTTGATATTTTCTAATAGAATTTTATGCTGTCTATCAGAAATACCATTTACAAA  
 TAATTTTAAACGCAAAAAAAGTTATAAATGTTTAGTGTAAACCTGGACACAAC  
 TACATAATGATTCTGATTTTGTAAAAAACAACAAAAACACACATATACACATGCA  
 TACATATGCATATAAAGAAACTGGAACAAACAAAATAACAAGCATAGTTGGAATTACAG  
 TCATTTTAATATCTTTATGCTTTTAAAAATTTGAAGTTGTATTACTAGCATCCACTA  
 CTTACGTAGTCAGGAAAAAATAACAACCTTAAATAGATATTTAGGTCCAAAGATGGTAA

FIGURE 15.4

37/64

TCTAAATGGTGTACAGGCTGAATGTGTGCCTGATCCCCATGCCCAAGTTCATATGTTA  
 AAGCCCTGGCCCCCAAGGCAATGGTATTAGGGGAGTAGGGCCTTTGGGAGGTAATCAGAT  
 TTCTACGAGGTATGAGGGTGGAGCCCGCATAGTGAATTAGTGTCTTTTAGGAAGAGG  
 AGAACAGACCAAAGCCTTCCTTTCTCTCCTCACTATGTAAGAAGACAGCCAGAAGGTGGC  
 CACAGCCAGGAAGAGAGCTCTCACCAGAACCCTGCTAGCACCTTGCTCTTGGGTT  
 CTCAGCATCCAGAAGTGTGAGAAATGAATGTGTGTTGTTTAAACCACTCAGGCTACGGTA  
 TTTTGTGTCAGCAGCCCAAGCTGACAGAGATAGAAACAACAAGGACCCATCAGCAGAC  
 GAATGGATGATCAAAACGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT  
 ATGAAATCTGATACATGCTATAATGATGAACCTTGAAACATGTTAATGGAAATAAGCC  
 AAACCTTAAAGGACAAATATTGTATAATTCCACTTATATGAGTTAGTTACCTAGAATAGG  
 CAAATTATGTACAGAACATTAGAGGTTACCAGGGTTGTGGGAAGAGGGGTATT  
 GTGGGTACAAATTTTCGGTTTGGAGTGATTTTGAAAAAATCTGGAAATGGGTAGTGACA  
 GTAGTCAACATGATGAATGTACTTAATGACACTAAATTGTACACTTAAAAATGGTTAATA  
 CTGGGCTGGCGCAGTGGCTCATGGCTGTAAATCCAGAACCTTTGGGAGGCCAAGACAGGC  
 GGATCATGAGGTGAGGAGATTGAGACCATCTGGCTAACATGGTGAACCCCTGTCTCTAC  
 TAAAAATAAAAACAAATAAAAAAAATAGCCGGGCATGGTGGCAGGCACCTGTAGTC  
 CCAGCTACTCGGAGGCTGAGGCAGGAGAATGGTGTGACCTGGGAGTCGGAGCTTGCACT  
 GAGCTGAGATCGCGCCACTGCCTCCAGCCTGGGCAACAGAGCCAGATTCCGTCTCAAAA  
 AAAAAAAGGTTGATACCTGGGTGCGGTGGCTCATGCCTGTAATTTGAGCACTTT  
 GGGAGGCCAAGGCAGGCAGATCAGTTGAGGTCAAGAGTTAAGGACCAGCCTGGCCAACGT  
 GGCGAAACCCCATCTCTATTAAAAATACAAAAATAGTCGAGTGTGGTGGTGGTGGTGGTGGT  
 TAGTCCCAGCTGCTGGGAGGATGAGGCCTAGGAATTGCTTGAACCCAGGAGGCAGAGGTT  
 GCAGTGAGTTGAGATTGCGCCACTGCCTCCAGCCTGGGGACAGAGCGAGACTTAGTCT  
 CAAAAAAGGTTAAATTTGTAAGTTTGTATGCATATTTACCATAATCTTAAAAAA  
 TAGATGATTAGGAGTAAAGTCAACAGAATTTAATAACCAGTTGTAAATAGAGACTGAGT  
 GAGGAGGATGAATTAAGGAAGACATTGAGTACAACCTTTTGGTAGGTGAAAAACTCTTAA  
 AAAAATACGTGGGCAAAGATCCTACTTGATTCTTATAATTTAAAAATCTCCAGTTAGTA  
 AACAAGGCTAGGTGGAGATTGTCATGTGATGTGAGGTGTGTGTTCTGTTTGTAAATGTGA  
 GGAATGTGAGCCATCTCCTGGACTTGAATATCCATTAGATAATTGAAATACGGATTGGA  
 GAACCTCAGGAGACGTGCAATGCAGTAACAAAACCTCTGCACCTAGTTGATTCTGTCTCCT  
 AATTTAATGCTTTATGGGACAACTGTTAGGCAGGTGGGCAAGATGGACAGCCATATTT  
 TTGTGGGTTCTTGGCCTGTGGCCAGCCTCAGTGCTCACTCTGAGGTGATGTCCAACTT  
 AGAACACATTAGGCCTACCACAGTCAAGGCTCCCTTTCTCAACTCTAGTCTCTGCACA  
 AATATCCGAAGCCTAGAAATAATAATCATCTGTCTTGTGCTTGCATTATGAAAGCCTA  
 GGAAAGGGCCTTGGGAATTAAGAAGAAATGGAAGAACTGGTCTAACTGCTGCATGCTTCAG  
 CTTGCAAGGGAACTCACTGAAATGGGGACAGGCCATAAAGGACAAACAGAGAGTGGCTT  
 CAGCAAGGCATCGTTTTTTCAGAGCAAGCTAGAGAATCCTGCCAGCGTCTCAGGCAGGG  
 CCCCTGGGCACAGAGGTTAGGCAAGGGAGTGTCCAGCATGTTGATGCCCTGAGCATCAG  
 AATAATGCCATAGAGGAGCTTCCAAAGAGTTCAATTCAGGTTTTGTAAGCCGAACATTTC  
 TAGGCAAAATAAAATTTGATTTGTGAATAAAGCTTGTCTTCAACTCCAGTGCAGATTC  
 TCATAGATTGATAGTGGCTTGTGATCCAGATAAAGAAAAACAATTTTCAAAGATTCTAT  
 TCTTTGTAGATGTACGGATTAGAGACCATCTAATCTAACTCCCTCATTCTACAGATAGG  
 AAAAAATGAGGCCATAAAGAGTTAAGAAAATACCATGGAAATGTCACTGCTGAACTGCCAT  
 ACGTAGGATCCGAAAGAAATTGGGTAAATGTACTGTGAGAAATACAGTACTAGGTCCAA  
 AGAATCTAATACAAATTAATAATCTAAATGTTATTTCTAAAGCATCCCTGCACATGGCTG  
 AACTTACATAGTTTCAATTTCTTTCTTTCTGTTGAAGAAGAGGCAATTGGCTGGGTGCA  
 GTGGCTCATGCCTGTAATCCTGGCACTTTGAGAGGCCGAGGCGGGTGGATCACCTGAGGT  
 CAGGAGTTTGAAGACAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAAATACAAA  
 AATTAGCTGGCTGTGGTGGCCGCTGCCTGTAATCCAGCTACTCCAGAGGCTGAGGCAGG  
 AGAATTACTTGAAATCTGGGAGGTGGAGGTTGAGTGAGCCAAGATCACGCCATTGCACTC  
 TAGCCTGGATGACAAGAGGGAACTCCATCTCAAAAAAAGAAAAAAGCAATCACT  
 AACCTGTGTTGTTTATTAAACATGACAGACTGGCATGAAGTAATTACCAACTGTAAACA

FIGURE 15.5

38/64

AAAAAGCTACAATCTGCCAGGCATGGTGGCTCATGCCTGTAATCCCCACCTTGGGAGGC  
CAGGTTGGGGGATCACCTGAGGCCTGGAGTTCAAGACTAGCCTGGTCAACATGGTGAAAC  
CTCGTCTCTACTAAAAATACAAAAATTAGCCCGCGTGGTGGCACATCCCTGTAATCCCA  
GTTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGCAGTGGGGAGGTTGCAGT  
GAGCCAAGATCGCACCCTTGTACTCCAGTCTGGGCCGACAGAGTGGAGTCTCGGTCTCAA  
AAAAAGAAAAAGAAAAAGCTACAACCTTAATCTCAACTTCTCATAACATCATCTCTACTT  
CTGATTAGAAGAGTGGAGTGGGGAGGTTTATTACAAAAAGACTGTTATACCTTACACAC  
TTCTCCCCATGAATAGTGAAGGTGTGAGTGAAAAAGACAGCAATTTTATTTTTTTTGA  
AACAGGTTCTTGCACTGTCAACCGGGCTGGAGTGCCTGTTGTGATCACTGCTCACTGCA  
GCCTCCACCTCCCAGGCTCAAGTGATCCTCCTACCTCAGCCTCCTGAGTAGCTGGGACCA  
CAGTTGTGCACTACCATGCCAGCTATTTTAAAAAGAGATGGGGTCTCACTATATTGC  
TTAGGCTAGTTCTCAAACTCCTGGCCTCAAGCAGTCTCCGACCTTGGCCTCCCAAAGGG  
TTGTGATTACAGGCATAAGCCACCACACCCAGCCAGCAGTTTATAGAATAAAGGGTGAAGG  
TGCTGTTGGGGAAATATAATTTAAAAAACAAAATCTTCTCTCAACCCAGAAATCCTCTCC  
ATGAAGGCAGTAGAGAAAGATAAGCTTTATTATTGAATAAAAAATTAAATGAGAATGTGAT  
GCACATCACAGGCACCTTTGCTAAGAGATCACAAAGACAGAAGGAAATTCACCATTGT  
ACAGCCAAGCAGGTACAGCCATTACATGTATGTTTTCGAGATAAATAGTCTCAACTAA  
GAGAACTTGACAGCACCCTGGTCACACAGTTCACTTAACCTTACCTGATAATTGATGT  
GACCACTTGTTATCTAAGATATCAACTTTTCGGGGGTGGGGGAGTGTGAAACAGGAG  
TTACTTTTATAGCTTGGTGCAAGGTACTCATTAGATTAGGCTGTTACCCTCCACAGAA  
ACTGGAAGATAGGTATGCTATCTGGTAATGTTTACATTTCCAGATCCTTGAGAAAGACA  
TTCTTAGGTCATAAAGCTGACAAAAGGCTGATTACAGTTTTTAAATATATATCTGTATA  
TGTATTCA

FIGURE 15.6



39/64

actgagagacaggactagctggatttcctaggctgactaagaatccctaagcctagctgg  
 |||||  
 actgagagacaggactagctggatttcccaggccgactaagaattcctaagcctagctgg  
  
 g-aagggtgaccacatccacctttaaacacggggcttgcaacttagctcacacctgaccaa  
 | |||||  
 ggaagggtgaccacacctcctttaaacacagagcttgtaactcagctcacacccgaccaa  
  
 tcag-----agagctcactaaaatgctaattaggc-aaagacaggaggtaaagaaa  
 |||||  
 tcaggtagtaagagagctcactaaaataccaattaggctaaaaacaggaggtaaagaaa  
  
 tagccaa-tcatctattgcctgagagcacagcaggagggaacaatgatcgggatataaac  
 || |||  
 taatcaaatcatctatcgctgagagcacagggggagggaacaatgatcgggatataaac  
  
 caagtcttcgagccggcaacggcaacccccctttgggtccctcctttgtatgggagctc  
 || | |||||  
 caggcatattgagccagatcaggtaaccctcctttgggtccctcacactgtatgggagctc  
  
 tgttttcatgctatttcactctattaaatcttgcaactgcac--tcttctgggtccatgtt  
 |||||  
 tgtt-----ttcactctattaaatcttgcaactgcacactcttctgggtccatgtt  
  
 tcttacggcttgagctgagctttcgctcgccatccaccactgctgtttgcccaccgca  
 | || |||||  
 tgttcggctcaagctgagcttttgctcgccgtccaccactgctgaatgcccattgca  
  
 gaccgcccgtgactcccatccctctggatcatgcagggtgtccgctgtgctcctgatcc  
 |||||  
 gacctgcccttgacttcacccctccggatccggcagagtgtccgctgactcctgatcc  
  
 agcgaggcacccattgccgctcccaatcgggctaaaggcttgccattgttcctgcatggc  
 |||||  
 agcgaggcacccattgccactcccgatcaggctaaaggcttgccattgttcctgcacagc  
  
 taagtgcctgggttcacctaattgagctgaacactagtactgggttccatgggttctct  
 |||||  
 taagtgcctgggttcacctaatacaggctgaacactgggtcgctgggttccacgggttctct  
  
 tctgtgaccacagcttctaatagagctataaactcaccgcatggcccaagggttccatt  
 || |||||  
 tccatgactcacagcttctaatagagctataaactcaccacatggcccaagggttccatt  
  
 cctt-gaatccataaggccaagaaccccagggtcagagaacacgaggcttgccaccatctt  
 | || |||||  
 cgttggaatccatgaggccaagaaccccagggtcagagaataaaaggcccgcc-ccatctt  
  
 gggag  
 |||||  
 gggag

FIGURE 16

## 40/64

TCCTGTGAAC CTCTAGAGGA TTTGCGCCTG CTCTTCAAAC AACCAACCAGG AGGAAAGTAA	7660
CTAAAATCAT AAATCCCAT GGCCTCCT TATCATATT TTCTCTTAC TGTCTTTA	7920
CCCTCTTTCA CTCTCACTGC ACCCCTCCA TGCCGCTGTA TGACCAGTAG CTCCCTTAC	7980
CAAGAGTTT TATGGAGAAT GCAGCGTCCC GGAAATATTG ATGCCCATC GTATAGGAGT	8040
CTTTCTAAGG GAACCCAC CTTCACTGCC CACACCCATA TGCCCGCAA CTGCTATCAC	8100
TCTGCCACTC TTTGCATGCA TGCAAACTACT CATTATTGGA CAGGAAAAAT GATTAATCCT	8160
AGTTGTCCTG GAGGACTTGG AGTCACTGTC TGTTGGACTT ACTTCACCCA AACTGGTATG	8220
TCTGATGGGG GTGGAGTTCA AGATCAGGCA AGAGAAAAAC ATGTAAAAGA AGTAATCTCC	8280
CAACTCACC GGGTACATGG CACCTCTAGC CCCTACAAAG GACTAGATCT CTCAAACTA	8340
CATGAAACCC TCCGTACCCA TACTCGCCTG GTAAGCCTAT TTAATACCAC CCTCACTGGG	8400
CTCCATGAGG TCTCGGCCA AAACCCTACT AACTGTTGGA TATGCCTCCC CCTGA <sup>u</sup> ACTTC	8460
AGGCCATAT <sup>u</sup> TTTCAATCCC TGTACCTGAA CAATGGAACA ACTTCAGCAC AGAAATAAAC	8520
ACCACTTCCG TTTTAGTAGG ACCTCTTGT TCCAATCTGG AAATAACCCA TACCTCAAAC	8580
CTCACCTGTG TAAATTTAG CAATACTACA TACACAACCA ACTCCCAATG CATCAGGTGG	8640
GTA <sup>u</sup> ACTCCTC CCACACA <sup>u</sup> AAT AGTCTGCCTA CCCTCAGGAA TATTTTTTGT CTGTGGTACC	8700
TCAGCCTATC GTTGT <sup>u</sup> TTGAA TGGCTCTTCA GAATCTATGT GCTTCCTCTC ATTCTTAGTG	8760
CCCCCTATGA CCATCTACAC TGAACAAGAT TTATACAGTT ATGTCATATC TAAGCCCCGC	8820
AACAAAAGAG TACCCATTCT TCCTTTTGT ATAGGAGCAG GAGTGCTAGG TGCACTAGGT	8880
ACTGGCATTG GCGGTATCAC AACCTCTACT CAGTTCTACT ACAAACTATC TCAAGAACTA	8940
AATGGGGACA TGGAACGGT CGCCGACTCC CTGGTCACCT TGCAAGATCA ACTTAACTCC	9000
CTAGCAGCAG TAGTCCTTCA AAATCGAAGA GCTTTAGACT TGCTAACCGC TGAAAGAGGG	9060
GGAACCTGTT TATTTTTAGG GGAAGAATGC TGTTATTATG TTAATCAATC CGGAATCGTC	9120
ACTGAGAAAG TTAAGAAAT TCGAGATCGA ATACAACGTA GAGCAGAGGA GCTTCGAAAC	9180
ACTGGACCTT GGGGCTCCT CAGCCAATGG ATGCCCTGGA TTCTCCCTT CTAGGACCT	9240
CTAGCAGCTA TAATATTGCT ACTCCTCTTT GGACCCTGTA TCTTTAACT CCTTGTTAAC	9300
TTTGTCTCTT CCAGAATCGA AGCTGTAAAA CTACAATGG AGCCCAAGAT GCAGTCCAAG	9360
ACTAAGATCT ACCGCAGACC CCTGGACCGG CCTGCTAGCC CACGATCTGA TGTTAATGAC	9420
ATCAAAGGCA CCCCTCTGA GGAATCTCA GCTGCACAAC CTCTACTAG CCCCATTCA	9480
GCAGGAAGCA GTTAGAGCGG TCTCGGCCA CCTCCCCAAC AGCACTTAGG TTTCTCTGTT	9540

FIGURE 17

41/64

AAGCTCCTTCAGGAGAACAAAGAACAGGCCATTACCCTGGAGAAGACTGGCAACTGATTTTACCCACAAGCCCAA  
 LysLeuLeuGlnGluAsnLysGluGlnAlaIleThrLeuGluLysThrGlyAsn...PheTyrProGlnAlaGln  
 SerSerPheArgArgThrLysAsnArgProLeuProTrpArgArgLeuAlaThrAspPheThrHisLysProLys  
 AlaProSerGlyGluGlnArgThrGlyHisTyrProGlyGluAspTrpGlnLeuIleLeuProThrSerProAsn

ACCTCAGGGATTTTCAGTATCTACTAGTCTGGGTAGATACTTTTCACGGGTTGGGCAGAGGCCTTCCCTGTAGGAC  
 ThrSerGlyIleSerValSerThrSerLeuGlyArgTyrPheHisGlyLeuGlyArgGlyLeuProLeu...Asp  
 ProGlnGlyPheGlnTyrLeuLeuValTrpValAspThrPheThrGlyTrpAlaGluAlaPheProCysArgThr  
 LeuArgAspPheSerIleTyr...SerGly...IleLeuSerArgValGlyGlnArgProSerProValGlyGln

AGAAAAGGCCCAAGAGGTAATAAAGGCACTAGTTCATGAAATAATCCCAGATTTCGGACTTCCCCGAGGCTTACA  
 ArgLysGlyProArgGlyAsnLysGlyThrSerSer...AsnAsnSerGlnIleArgThrSerProArgLeuThr  
 GluLysAlaGlnGluValIleLysAlaLeuValHisGluIleIleProArgPheGlyLeuProArgGlyLeuGln  
 LysArgProLysArg.....ArgHis...PheMETLys...PheProAspSerAspPheProGluAlaTyrArg

GAGTGACAATAGCCCTGCTTTCCAGGCCACAGTAACCCAGGGAGTATCCCAGGCGTTAGGTATACGATATCACTT  
 Glu...Gln...ProCysPheProGlyHisSerAsnProGlySerIleProGlyValArgTyrThrIleSerLeu  
 SerAspAsnSerProAlaPheGlnAlaThrValThrGlnGlyValSerGlnAlaLeuGlyIleArgTyrHisLeu  
 ValThrIleAlaLeuLeuSerArgProGln...ProArgGluTyrProArgArg...ValTyrAspIleThrTyr

ACACTGCGCCTGAAGGCCACAGTCCTCAGGAAGGTCGAGAAAATGAATGAAACACTCAAAGGACATCTAAAAAA  
 ThrLeuArgLeuLysAlaThrValLeuArgGluGlyArgGluAsnGlu...AsnThrGlnArgThrSerLysLys  
 HisCysAla...ArgProGlnSerSerGlyLysValGluLysMETAsnGluThrLeuLysGlyHisLeuLysLys  
 ThrAlaProGluGlyHisSerProGlnGlyArgSerArgLys...METLysHisSerLysAspIle...LysSer

GCAAACCCAGGAAACCCACCTCACATGGCCTGCTCTGTTGCCTATAGCCTTAAAAAGAATCTGCAACTTTCCCCA  
 385            395            405            415            425            435            445  
 AlaAsnProGlyAsnProProHisMETAlaCysSerValAlaTyrSerLeuLysLysAsnLeuGlnLeuSerPro  
 GlnThrGlnGluThrHisLeuThrTrpProAlaLeuLeuProIleAlaLeuLysArgIleCysAsnPheProGln  
 LysProArgLysProThrSerHisGlyLeuLeuCysCysLeu...Pro...LysGluSerAlaThrPheProLys

AAAAGCAGGACTTAGCCCATACGAAATGCTGTATGGAAGGCCCTTCATAACCAATGACCTTGTGCTTGACCCAAG  
 LysSerArgThr...ProIleArgAsnAlaValTrpLysAlaLeuHisAsnGln...ProCysAla...ProLys  
 LysAlaGlyLeuSerProTyrGluMETLeuTyrGlyArgProPheIleThrAsnAspLeuValLeuAspProArg  
 LysGlnAspLeuAlaHisThrLysCysCysMETGluGlyProSer...ProMETThrLeuCysLeuThrGlnAsp

ACAGCCAACTTAGTTGCAGACATCACCTCCTTAGCCAAATATCAACAAGTTCTTAAACATTACAAGGAACCTAT  
 ThrAlaAsnLeuValAlaAspIleThrSerLeuAlaLysTyrGlnGlnValLeuLysThrLeuGlnGlyThrTyr  
 GlnProThr...LeuGlnThrSerProPro...ProAsnIleAsnLysPheLeuLysHisTyrLysGluProIle  
 SerGlnLeuSerCysArgHisHisLeuLeuSerGlnIleSerThrSerSer...AsnIleThrArgAsnLeuSer

CCCTGAGAAGAGGGAAAAGAACTATTCCACCCTTGTGACATGGTATTAGTCAAGTCCCTTCCCTCTAATTCCCCA  
 Pro...GluGluGlyLysGluLeuPheHisProCysAspMETValLeuValLysSerLeuProSerAsnSerPro  
 ProGluLysArgGluLysAsnTyrSerThrLeuValThrTrpTyr...SerSerProPheProLeuIleProHis  
 LeuArgArgGlyLysArgThrIleProProLeu...HisGlyIleSerGlnValProSerLeu...PheProIle

TCCCTAGATACATCTGGGAAGGACCCTACCCAGTCATTTTATCTACCCCAACTGCGGTTAAAGTGGCTGGAGTG  
 SerLeuAspThrSerTrpGluGlyProTyrProValIleLeuSerThrProThrAlaValLysValAlaGlyVal  
 Pro...IleHisProGlyLysAspProThrGlnSerPheTyrLeuProGlnLeuArgLeuLysTrpLeuGluTrp  
 ProArgTyrIleLeuGlyArgThrLeuProSerHisPheIleTyrProAsnCysGly...SerGlyTrpSerGly

FIGURE 18.1

42/64

GAGTCTTGGATACATCACAACCTTGAGTCAAATCCTGGATAGTCCCAAAGGAACCTGAAAATCCAGGAGACAACGCT  
 GluSerTrpIleHisHisThr...ValLysSerTrpIleLeuProLysGluProGluAsnProGlyAspAsnAla  
 SerLeuGlyTyrIleThrLeuGluSerAsnProGlyTyrCysGlnArgAsnLeuLysIleGlnGluThrThrLeu  
 ValLeuAspThrSerHisLeuSerGlnIleLeuAspThrAlaLysGlyThr...LysSerArgArgGlnArg...

AGCTATTCTCTGTAACCTCTAGAGGATTTGCGCCTGCTCTTCAAACAACAACCAGGAGGAAAGTAACTAAAATCA  
 SerTyrSerCysGluProLeuGluAspLeuArgLeuLeuPheLysGlnGlnProGlyGlyLys...LeuLysSer  
 AlaIleProValAsnLeu...ArgIleCysAlaCysSerSerAsnAsnAsnGlnGluGluSerAsn...AsnHis  
 LeuPheLeu...ThrSerArgGlyPheAlaProAlaLeuGlnThrThrThrArgArgLysValThrLysIleIle

TAAATCCCCATGGCCCTCCCTTATCATATTTTCTCTTTACTGTTCTTTACCCTCTTTCACTCTCACTGCACCC  
 ...IleProMETAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThrAlaPro  
 LysSerProTrpProSerLeuIleIlePhePheSerLeuLeuPhePheTyrProLeuSerLeuSerLeuHisPro  
 AsnProHisGlyProProLeuSerTyrPheSerLeuTyrCysSerPheThrLeuPheHisSerHisCysThrPro

CCTCCATGCCGCTGTATGACCAGTAGCTCCCTTACCAAGAGTTTCTATGGAGAATGCAGCGTCCCGGAAATATT  
ProProCysArgCysMETThrSerSerSerProTyrGlnGluPheLeuTrpArgMETGlnArgProGlyAsnIle  
 LeuHisAlaAlaVal...ProValAlaProLeuThrLysSerPheTyrGlyGluCysSerValProGluIleLeu  
 SerMETProLeuTyrAspGln...LeuProLeuProArgValSerMETGluAsnAlaAlaSerArgLysTyr...

GATGCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCCACCTTCACTGCCCCACCCCATATGCCCGCAACTGC  
AspAlaProSerTyrArgSerLeuSerLysGlyThrProThrPheThrAlaHisThrHisMETProArgAsnCys  
 METProHisArgIleGlyValPheLeuArgGluProProProSerLeuProThrProIleCysProAlaThrAla  
 CysProIleVal...GluSerPhe...GlyAsnProHisLeuHisCysProHisProTyrAlaProGlnLeuLeu

TATCACTCTGCCACTCTTTGCATGCATGCAAATACTCATTATTGGACAGGAAAAATGATTAATCCTAGTTGTCTCT  
TyrHisSerAlaThrLeuCysMETHisAlaAsnThrHisTyrTrpThrGlyLysMETIleAsnProSerCysPro  
 IleThrLeuProLeuPheAlaCysMETGlnIleLeuIleIleGlyGlnGluLys...LeuIleLeuValValLeu  
 SerLeuCysHisSerLeuHisAlaCysLysTyrSerLeuLeuAspArgLysAsnAsp...Ser...LeuSerTrp

GGAGGACTTGGAGTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGTCTGATGGGGGTGGAGTTCAAGAT  
GlyGlyLeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMETSerAspGlyGlyGlyValGlnAsp  
 GluAspLeuGluSerLeuSerValGlyLeuThrSerProLysLeuValCysLeuMETGlyValGluPheLysIle  
 ArgThrTrpSerHisCysLeuLeuAspLeuLeuHisProAsnTrpTyrVal...TrpGlyTrpSerSerArgSer

CAGGCAAGAGAAAAACATGTAAAAGAAGTAATCTCCCAACTCACCCGGGTACATGGCACCTCTAGCCCCCTACAAA  
GlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHisGlyThrSerSerProTyrLys  
 ArgGlnGluLysAsnMET...LysLys...SerProAsnSerProGlyTyrMETAlaProLeuAlaProThrLys  
 GlyLysArgLysThrCysLysArgSerAsnLeuProThrHisProGlyThrTrpHisLeu...ProLeuGlnArg

GGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACCCATACTCGCCTGGTAAGCCTATTTAATACCACCCTC  
GlyLeuAspLeuSerLysLeuHisGluThrLeuArgThrHisThrArgLeuValSerLeuPheAsnThrThrLeu  
 Asp...IleSerGlnAsnTyrMETLysProSerValProIleLeuAlaTrp...AlaTyrLeuIleProProSer  
 ThrArgSerLeuLysThrThr...AsnProProTyrProTyrSerProGlyLysProIle...TyrHisProHis

ACTGGGCTCCATGAGGTCTCGGCCCAAAACCTACTAACTGTTGGATATGCCTCCCCCTGAACTTCAGGCCATAT  
ThrGlyLeuHisGluValSerAlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyr  
 LeuGlySerMETArgSerArgProLysThrLeuLeuThrValGlyTyrAlaSerPro...ThrSerGlyHisMET  
 TrpAlaPro...GlyLeuGlyProLysProTyr...LeuLeuAspMETProProProGluLeuGlnAlaIleCys

GTTTCATCCCTGTACCTGAACAATGGAACAACCTCAGCACAGAAATAAACACCACTTCCGTTTTAGTAGGACCT  
ValSerIleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyPro  
 PheGlnSerLeuTyrLeuAsnAsnGlyThrThrSerAlaGlnLys...ThrProLeuProPhe.....AspLeu  
 PheAsnProCysThr...ThrMETGluGlnLeuGlnHisArgAsnLysHisHisPheArgPheSerArgThrSer

FIGURE 18.2

43/64

CTTGTTTCCAATCTGGAAATAACCCATACCTCAAACCTCACCTGTGTAAATTTAGCAATACTACATACACAACC  
LeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrThrTyrThrThr  
 LeuPheProIleTrpLys...ProIleProGlnThrSerProVal...AsnLeuAlaIleLeuHisThrGlnPro  
 CysPheGlnSerGlyAsnAsnProTyrLeuLysProHisLeuCysLysIle...GlnTyrTyrIleHisAsnGln

AACTCCCAATGCATCAGGTGGGTAACTCCTCCACACAAATAGTCTGCCTACCCTCAGGAATATTTTTGTCTGT  
AsnSerGlnCysIleArgTrpValThrProProThrGlnIleValCysLeuProSerGlvIlePhePheValCys  
 ThrProAsnAlaSerGlyGly...LeuLeuProHisLys...SerAlaTyrProGlnGluTyrPheLeuSerVal  
 LeuProMETHisGlnValGlyAsnSerSerHisThrAsnSerLeuProThrLeuArgAsnIlePheCysLeuTrp

GGTACCTCAGCTATCGTTGTTTGAATGGCTCTTCAGAATCTATGTGCTTCCTCTCATTCTTAGTGCCCCCTATG  
GlyThrSerAlaTyrArgCysLeuAsnGlySerSerGluSerMETCysPheLeuSerPheLeuValProProMET  
 ValProGlnProIleValVal...METAlaLeuGlnAsnLeuCysAlaSerSerHisSer...CysProLeu...  
 TyrLeuSerLeuSerLeuPheGluTrpLeuPheArgIleTyrValLeuProLeuIleLeuSerAlaProTyrAsp

ACCATCTACACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCCGCAACAAAAGAGTACCCATTCTTCCT  
ThrIleTyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIleLeuPro  
 ProSerThrLeuAsnLysIleTyrThrValMETSerTyrLeuSerProAlaThrLysGluTyrProPhePheLeu  
 HisLeuHis...ThrArgPheIleGlnLeuCysHisIle...AlaProGlnGlnLysSerThrHisSerSerPhe

TTTGTTATAGGAGCAGGAGTGCTAGGTGCACTAGGTACTGGCATTGGCGGTATCACAACCTCTACTCAGTTCTAC  
PheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIleThrThrSerThrGlnPheTyr  
 LeuLeu...GluGlnGluCys...ValHis...ValLeuAlaLeuAlaValSerGlnProLeuLeuSerSerThr  
 CysTyrArgSerArgSerAlaArgCysThrArgTyrTrpHisTrpArgTyrHisAsnLeuTyrSerValLeuLeu

TACAAACTATCTCAAGAACTAAATGGGGACATGGAACGGGTCGCCGACTCCCTGGTCACCTTGCAAGATCAACTT  
TyrLysLeuSerGlnGluLeuAsnGlyAspMETGluArgValAlaAspSerLeuValThrLeuGlnAspGlnLeu  
 ThrAsnTyrLeuLysAsn...METGlyThrTrpAsnGlySerProThrProTrpSerProCysLysIleAsnLeu  
 GlnThrIleSerArgThrLysTrpGlyHisGlyThrGlyArgArgLeuProGlyHisLeuAlaArgSerThr...

AACTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGAACCTGT  
AsnSerLeuAlaAlaValValLeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCys  
 ThrPro...GlnGln...SerPheLysIleGluGluLeu...ThrCys...ProLeuLysGluGlyGluProVal  
 LeuProSerSerSerSerProSerLysSerLysSerPheArgLeuAlaAsnArg...LysArgGlyAsnLeuPhe

TTATTTTATAGGGGAAGAATGCTGTTATTATGTTAATCAATCCGAATCGTCACTGAGAAAGTTAAAGAAATTCGA  
LeuPheLeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGluIleArg  
 TyrPhe...GlyLysAsnAlaValIleMETLeuIleAsnProGluSerSerLeuArgLysLeuLysLysPheGlu  
 IlePheArgGlyArgMETLeuLeuLeuCys...SerIleArgAsnArgHis...GluSer...ArgAsnSerArg

GATCGAATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCCTGGGGCCTCCTCAGCCAATGGATGCCCTGG  
AspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeuLeuSerGlnTrpMETProTrp  
 IleGluTyrAsnValGluGlnArgSerPheGluThrLeuAspProGlyAlaSerSerAlaAsnGlyCysProGly  
 SerAsnThrThr...SerArgGlyAlaSerLysHisTrpThrLeuGlyProProGlnProMETAspAlaLeuAsp

ATTCTCCCCTTCTTAGGACCTCTAGCAGCTATAATATTGCTACTCCTCTTTGGACCCTGTATCTTTAACCTCCTT  
IleLeuProPheLeuGlyProLeuAlaAlaIleIleLeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeu  
 PheSerProSer...AspLeu...GlnLeu...TyrCysTyrSerSerLeuAspProValSerLeuThrSerLeu  
 SerProLeuLeuArgThrSerSerSerTyrAsnIleAlaThrProLeuTrpThrLeuTyrLeu...ProProCys

FIGURE 18.3

44/64

GTTAAC TTTGTCTCTTCCAGAATCGAAGCTGTAAACTACAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATC  
ValAsnPheValSerSerArgIleGluAlaValLysLeuGlnMETGluProLysMETGlnSerLysThrLysIle  
LeuThrLeuSerLeuProGluSerLysLeu...AsnTyrLysTrpSerProArgCysSerProArgLeuArgSer  
...LeuCysLeuPheGlnAsnArgSerCysLysThrThrAsnGlyAlaGlnAspAlaValGlnAsp...AspLeu

TACCGCAGACCCCTGGACCGGCTGCTAGCCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCTGAGGAA  
TyrArgArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProProGluGlu  
ThrAlaAspProTrpThrGlyLeuLeuAlaHisAspLeuMETLeuMETThrSerLysAlaProLeuLeuArgLys  
ProGlnThrProGlyProAlaCys...ProThrIle...Cys.....HisGlnArgHisProSer...GlyAsn

ATCTCAGCTGCACAACCTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGTCGTCGGCCAACCTCCCCA  
IleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer...SerGlyArgArgProThrSerPro  
SerGlnLeuHisAsnLeuTyrTyrAlaProIleGlnGlnGluAlaValArgAlaValValGlyGlnProProGln  
LeuSerCysThrThrSerThrThrProGlnPheSerArgLysGlnLeuGluArgSerSerAlaAsnLeuProAsn

ACAGCACTTAGGTTTCTCTGTTGAGATGGGGG  
ThrAlaLeuArgPheSerCys...AspGlyGly  
GlnHisLeuGlyPheProValGluMETGly  
SerThr...ValPheLeuLeuArgTrpGly

**FIGURE 18.4**

## 45/64

LysLeuLeuGlnGluAsnLysGluGlnAlaIleThrLeuGluLysThrGlyAsn...PheTyrProGlnAlaGln  
 ThrSerGlyIleSerValSerThrSerLeuGlyArgTyrPheHisGlyLeuGlyArgGlyLeuProLeu...Asp  
 ArgLysGlyProArgGlyAsnLysGlyThrSerSer...AsnAsnSerGlnIleArgThrSerProArgLeuThr  
 Glu...Gln...ProCysPheProGlyHisSerAsnProGlySerIleProGlyValArgTyrThrIleSerLeu  
 ThrLeuArgLeuLysAlaThrValLeuArgGluGlyArgGluAsnGlu...AsnThrGlnArgThrSerLysLys  
  
 AlaAsnProGlyAsnProProHisMETAlaCysSerValAlaTyrSerLeuLysLysAsnLeuGlnLeuSerPro  
 LysSerArgThr...ProIleArgAsnAlaValTrpLysAlaLeuHisAsnGln...ProCysAla...ProLys  
 ThrAlaAsnLeuValAlaAspIleThrSerLeuAlaLysTyrGlnGlnValLeuLysThrLeuGlnGlyThrTyr  
 Pro...GluGluGlyLysGluLeuPheHisProCysAspMETValLeuValLysSerLeuProSerAsnSerPro  
 SerLeuAspThrSerTrpGluGlyProTyrProValIleLeuSerThrProThrAlaValLysValAlaGlyVal  
 GluSerTrpIleHisHisThr...ValLysSerTrpIleLeuProLysGluProGluAsnProGlyAspAsnAla  
 SerTyrSerCysGluProLeuGluAspLeuArgLeuLeuPheLysGlnGlnProGlyGlyLys...LeuLysSer  
 ...IleProMETAlaLeuProTyrHisIlePheLeuPheThrValLeuLeuProSerPheThrLeuThrAlaPro  
 ProProCysArgCysMETThrSerSerSerProTyrGlnGluPheLeuTrpArgMETGlnArgProGlyAsnIle  
 AspAlaProSerTyrArgSerLeuSerLysGlyThrProThrPheThrAlaHisThrHisMETProArgAsnCys  
 TyrHisSerAlaThrLeuCysMETHisAlaAsnThrHisTyrTrpThrGlyLysMETIleAsnProSerCysPro  
 GlyGlyLeuGlyValThrValCysTrpThrTyrPheThrGlnThrGlyMETSerAspGlyGlyGlyValGlnAsp  
 GlnAlaArgGluLysHisValLysGluValIleSerGlnLeuThrArgValHisGlyThrSerSerProTyrLys  
 GlyLeuAspLeuSerLysLeuHisGluThrLeuArgThrHisThrArgLeuValSerLeuPheAsnThrThrLeu  
 ThrGlyLeuHisGluValSerAlaGlnAsnProThrAsnCysTrpIleCysLeuProLeuAsnPheArgProTyr  
 ValSerIleProValProGluGlnTrpAsnAsnPheSerThrGluIleAsnThrThrSerValLeuValGlyPro  
 LeuValSerAsnLeuGluIleThrHisThrSerAsnLeuThrCysValLysPheSerAsnThrThrTyrThrThr  
 AsnSerGlnCysIleArgTrpValThrProProThrGlnIleValCysLeuProSerGlyIlePhePheValCys  
 GlyThrSerAlaTyrArgCysLeuAsnGlySerSerGluSerMETCysPheLeuSerPheLeuValProProMET  
 ThrIleTyrThrGluGlnAspLeuTyrSerTyrValIleSerLysProArgAsnLysArgValProIleLeuPro  
 PheValIleGlyAlaGlyValLeuGlyAlaLeuGlyThrGlyIleGlyGlyIleThrThrSerThrGlnPheTyr  
 TyrLysLeuSerGlnGluLeuAsnGlyAspMETGluArgValAlaAspSerLeuValThrLeuGlnAspGlnLeu

**FIGURE 19.1**

**46/64**

AsnSerLeuAlaAlaValValLeuGlnAsnArgArgAlaLeuAspLeuLeuThrAlaGluArgGlyGlyThrCys  
LeuPheLeuGlyGluGluCysCysTyrTyrValAsnGlnSerGlyIleValThrGluLysValLysGluIleArg  
AspArgIleGlnArgArgAlaGluGluLeuArgAsnThrGlyProTrpGlyLeuLeuSerGlnTrpMETProTrp  
IleLeuProPheLeuGlyProLeuAlaAlaIleIleLeuLeuLeuLeuPheGlyProCysIlePheAsnLeuLeu  
ValAsnPheValSerSerArgIleGluAlaValLysLeuGlnMETGluProLysMETGlnSerLysThrLysIle  
TyrArgArgProLeuAspArgProAlaSerProArgSerAspValAsnAspIleLysGlyThrProProGluGlu  
IleSerAlaAlaGlnProLeuLeuArgProAsnSerAlaGlySerSer...SerGlyArgArgProThrSerPro  
ThrAlaLeuArgPheSerCys...AspGlyGly

**FIGURE 19.2**



## 47/64

SerSerPheArgArgThrLysAsnArgProLeuProTrpArgArgLeuAlaThrAspPheThrHisLysProLys  
ProGlnGlyPheGlnTyrLeuLeuValTrpValAspThrPheThrGlyTrpAlaGluAlaPheProCysArgThr  
GluLysAlaGlnGluValIleLysAlaLeuValHisGluIleIleProArgPheGlyLeuProArgGlyLeuGln  
SerAspAsnSerProAlaPheGlnAlaThrValThrGlnGlyValSerGlnAlaLeuGlyIleArgTyrHisLeu  
HisCysAla...ArgProGlnSerSerGlyLysValGluLysMETAsnGluThrLeuLysGlyHisLeuLysLys  
GlnThrGlnGluThrHisLeuThrTrpProAlaLeuLeuProIleAlaLeuLysArgIleCysAsnPheProGln  
LysAlaGlyLeuSerProTyrGluMETLeuTyrGlyArgProPheIleThrAsnAspLeuValLeuAspProArg  
GlnProThr...LeuGlnThrSerProPro...ProAsnIleAsnLysPheLeuLysHisTyrLysGluProIle  
ProGluLysArgGluLysAsnTyrSerThrLeuValThrTrpTyr...SerSerProPheProLeuIleProHis  
Pro...IleHisProGlyLysAspProThrGlnSerPheTyrLeuProGlnLeuArgLeuLysTrpLeuGluTrp  
SerLeuGlyTyrIleThrLeuGluSerAsnProGlyTyrCysGlnArgAsnLeuLysIleGlnGluThrThrLeu  
AlaIleProValAsnLeu...ArgIleCysAlaCysSerSerAsnAsnAsnGlnGluGluSerAsn...AsnHis  
LysSerProTrpProSerLeuIleIlePhePheSerLeuLeuPhePheTyrProLeuSerLeuSerLeuHisPro  
LeuHisAlaAlaVal...ProValAlaProLeuThrLysSerPheTyrGlyGluCysSerValProGluIleLeu  
METProHisArgIleGlyValPheLeuArgGluProProProSerLeuProThrProIleCysProAlaThrAla  
IleThrLeuProLeuPheAlaCysMETGlnIleLeuIleIleGlyGlnGluLys...LeuIleLeuValValLeu  
GluAspLeuGluSerLeuSerValGlyLeuThrSerProLysLeuValCysLeuMETGlyValGluPheLysIle  
ArgGlnGluLysAsnMET...LysLys...SerProAsnSerProGlyTyrMETAlaProLeuAlaProThrLys  
Asp...IleSerGlnAsnTyrMETLysProSerValProIleLeuAlaTrp...AlaTyrLeuIleProProSer  
LeuGlySerMETArgSerArgProLysThrLeuLeuThrValGlyTyrAlaSerPro...ThrSerGlyHisMET  
PheGlnSerLeuTyrLeuAsnAsnGlyThrThrSerAlaGlnLys...ThrProLeuProPhe.....AspLeu  
LeuPheProIleTrpLys...ProIleProGlnThrSerProVal...AsnLeuAlaIleLeuHisThrGlnPro  
ThrProAsnAlaSerGlyGly...LeuLeuProHisLys...SerAlaTyrProGlnGluTyrPheLeuSerVal  
ValProGlnProIleValVal...METAlaLeuGlnAsnLeuCysAlaSerSerHisSer...CysProLeu...  
ProSerThrLeuAsnLysIleTyrThrValMETSerTyrLeuSerProAlaThrLysGluTyrProPhePheLeu  
LeuLeu...GluGlnGluCys...ValHis...ValLeuAlaLeuAlaValSerGlnProLeuLeuSerSerThr  
ThrAsnTyrLeuLysAsn...METGlyThrTrpAsnGlySerProThrProTrpSerProCysLysIleAsnLeu  
ThrPro...GlnGln...SerPheLysIleGluGluLeu...ThrCys...ProLeuLysGluGlyGluProVal

**FIGURE 20.1**

**48/64**

TyrPhe...GlyLysAsnAlaValIleMETLeuIleAsnProGluSerSerLeuArgLysLeuLysLysPheGlu  
IleGluTyrAsnValGluGlnArgSerPheGluThrLeuAspProGlyAlaSerSerAlaAsnGlyCysProGly  
PheSerProSer...AspLeu...GlnLeu...TyrCysTyrSerSerLeuAspProValSerLeuThrSerLeu  
LeuThrLeuSerLeuProGluSerLysLeu...AsnTyrLysTrpSerProArgCysSerProArgLeuArgSer  
ThrAlaAspProTrpThrGlyLeuLeuAlaHisAspLeuMETLeuMETThrSerLysAlaProLeuLeuArgLys  
SerGlnLeuHisAsnLeuTyrTyrAlaProIleGlnGlnGluAlaValArgAlaValValGlyGlnProProGln  
GlnHisLeuGlyPheProValGluMETGly

**FIGURE 20.2**

49/64

AlaProSerGlyGluGlnArgThrGlyHisTyrProGlyGluAspTrpGlnLeuIleLeuProThrSerProAsn  
 LeuArgAspPheSerIleTyr...SerGly...IleLeuSerArgValGlyGlnArgProSerProValGlyGln  
 LysArgProLysArg.....ArgHis...PheMETLys...PheProAspSerAspPheProGluAlaTyrArg  
 ValThrIleAlaLeuLeuSerArgProGln...ProArgGluTyrProArgArg...ValTyrAspIleThrTyr  
 ThrAlaProGluGlyHisSerProGlnGlyArgSerArgLys...METLysHisSerLysAspIle...LysSer  
 LysProArgLysProThrSerHisGlyLeuLeuCysCysLeu...Pro...LysGluSerAlaThrPheProLys  
 LysGlnAspLeuAlaHisThrLysCysCysMETGluGlyProSer...ProMETThrLeuCysLeuThrGlnAsp  
 SerGlnLeuSerCysArgHisHisLeuLeuSerGlnIleSerThrSerSer...AsnIleThrArgAsnLeuSer  
 LeuArgArgGlyLysArgThrIleProProLeu...HisGlyIleSerGlnValProSerLeu...PheProIle  
 ProArgTyrIleLeuGlyArgThrLeuProSerHisPheIleTyrProAsnCysGly...SerGlyTrpSerGly  
 ValLeuAspThrSerHisLeuSerGlnIleLeuAspThrAlaLysGlyThr...LysSerArgArgGlnArg...  
 LeuPheLeu...ThrSerArgGlyPheAlaProAlaLeuGlnThrThrThrArgArgLysValThrLysIleIle  
 AsnProHisGlyProProLeuSerTyrPheSerLeuTyrCysSerPheThrLeuPheHisSerHisCysThrPro  
 SerMETProLeuTyrAspGln...LeuProLeuProArgValSerMETGluAsnAlaAlaSerArgLysTyr...  
 CysProIleVal...GluSerPhe...GlyAsnProHisLeuHisCysProHisProTyrAlaProGlnLeuLeu  
 SerLeuCysHisSerLeuHisAlaCysLysTyrSerLeuLeuAspArgLysAsnAsp...Ser...LeuSerTrp  
 ArgThrTrpSerHisCysLeuLeuAspLeuLeuHisProAsnTrpTyrVal...TrpGlyTrpSerSerArgSer  
 GlyLysArgLysThrCysLysArgSerAsnLeuProThrHisProGlyThrTrpHisLeu...ProLeuGlnArg  
 ThrArgSerLeuLysThrThr...AsnProProTyrProTyrSerProGlyLysProIle...TyrHisProHis  
 TrpAlaPro...GlyLeuGlyProLysProTyr...LeuLeuAspMETProProProGluLeuGlnAlaIleCys  
 PheAsnProCysThr...ThrMETGluGlnLeuGlnHisArgAsnLysHisHisPheArgPheSerArgThrSer  
 CysPheGlnSerGlyAsnAsnProTyrLeuLysProHisLeuCysLysIle...GlnTyrTyrIleHisAsnGln  
 LeuProMETHisGlnValGlyAsnSerSerHisThrAsnSerLeuProThrLeuArgAsnIlePheCysLeuTrp  
 TyrLeuSerLeuSerLeuPheGluTrpLeuPheArgIleTyrValLeuProLeuIleLeuSerAlaProTyrAsp  
 HisLeuHis...ThrArgPheIleGlnLeuCysHisIle...AlaProGlnGlnLysSerThrHisSerSerPhe  
 CysTyrArgSerArgSerAlaArgCysThrArgTyrTrpHisTrpArgTyrHisAsnLeuTyrSerValLeuLeu  
 GlnThrIleSerArgThrLysTrpGlyHisGlyThrGlyArgArgLeuProGlyHisLeuAlaArgSerThr...  
 LeuProSerSerSerSerProSerLysSerLysSerPheArgLeuAlaAsnArg...LysArgGlyAsnLeuPhe

**FIGURE 21.1**

**50/64**

IlePheArgGlyArgMETLeuLeuLeuCys...SerIleArgAsnArgHis...GluSer...ArgAsnSerArg  
SerAsnThrThr...SerArgGlyAlaSerLysHisTrpThrLeuGlyProProGlnProMETAspAlaLeuAsp  
SerProLeuLeuArgThrSerSerSerTyrAsnIleAlaThrProLeuTrpThrLeuTyrLeu...ProProCys  
...LeuCysLeuPheGlnAsnArgSerCysLysThrThrAsnGlyAlaGlnAspAlaValGlnAsp...AspLeu  
ProGlnThrProGlyProAlaCys...ProThrIle...Cys.....HisGlnArgHisProSer...GlyAsn  
LeuSerCysThrThrSerThrThrProGlnPheSerArgLysGlnLeuGluArgSerSerAlaAsnLeuProAsn  
SerThr...ValPheLeuLeuArgTrpGly

**FIGURE 21.2**

51/64

TTGGTCTTAAGAACAACAAATGATATGGCTCCAATGACTGGAGGAACACCAGGGTCCTTGG  
TCTCACGCTGATTTAGATAAAACGACTGTTCAGGCCTCTGAGCCCAAGCTAAGCCATCCTC  
CCCTGTGACCTGCACGTATACATCCAGATGGCCTGAAGTAACCAAAGAATCACAAAAGCA  
GTGAAAATGGCCTGTTCTGCTTAACCTGATGACATTCACACCATTGTGATTTGTTCTGCTG  
CCCATCTTAACCTGAGCGATTAACTTGTGAAATTCCTTCTCCTGGCTCAAAACCTCCCC  
ACTGAGCACCTTGTGACCCCCGCCCCCTGCCCCCTAAGAGAAAACCCCTTTGATTATAATT  
TTCCACTACCCACCCAAATCCTATAAAATGGCCCCACCCCTATCTCCCTTCGCTGACTCC  
TTTTTCGGACTCAGCCCCGCTGCACCCAGGTGAAATAAACAGCCTTGTTGCTCACACAAA  
GCCTGTTTGGTGGACTCTCTTCACACGGACGCTCATGACATTTGGTGCCAAAACCTGGGA  
TAGGAGGACTCCTTCAGGAGACCAGTCCCCCTGTCCTTGCCCTCACTCTGTGAGGACATCC  
ACCTACAACCTTGGGTCTCAGACCAACCAGCCCAAGGAACAGCTCACCAATTTCAAATC  
AGGTAAGCAGTCTTTTCACTCTCTTCTCCAGCCTCTCTTGCTACCCTTCAAACCTCCCTCT  
CTCACTACCCTTCAATCTCCCTGTCTTCCAATTCCAGTCTTTTTTTCATCTCTAGTAGAG  
ACAAAGGAGACACATTTTATCCATGGACCCAAACTCCAGCACCAGTCACGGACTTGGGA  
AGACAGTCTTCCCTTGGTGTTAATCACTGCGGGGACGCTGCCTGATTATTCACCCACA  
CTCCATTGGTGTCTGATCAGGTGGGGACACCTGCCTTGGTCACTCACCCACATTCCCTT  
GGTGGTACGTCAACTGCAAAAGCAGGGGACGCCTGCTTGGCTGCTCACCCACCCCTTC  
TCTGTGTCTCTACCTTTCTCTTTAACTTACCTCCTTCACTATGGGCAAACCTTCTGCCCT  
CCATTCCTCCCTTCTTCTCCCTTAGCCTGTGTTCTTAAAAACCTAAACCTCTTCAACTCA  
CACCTGACCTAAACCTAAATGCCCTTATTTTCTTCTGCAACACTGCGTGGCTGCAGTACA  
AACTTGATAATAGCTTTAAATGGCCAGAATATGGCACTTCAATTTCTCCATCCTACAAG  
ATCTAGATAATTTTGTGGAATAATGGAAATGGTCTGAGATGCCTGACGTCCAGGCAT  
TCTTTTACACATTTGGTCCCTCCCTAGTCTCTGCTCCCAATGCGACTCATCCCAATCTTT  
CTTCTTTTCTCTCCTGTCTGTTCTTCACTCTCCACCCCAAGCTCTGAGTCTTTGAATCC  
TCCTTTGCTACAGACCCATCTGAACTCTCCCTCCTCCCCAGGCTGCTCCTCACCAGGCC  
GAGCCAGGTCCCAATTCTTCTCAGCCTCTGCTCCCCCACCCTATAATCCTTTTATCACC  
TCCTCTCCTCACACTCAGTCCGGCTTACAGTTTCGTTCTGTGACTAGCCCTCCCCATCT  
GCCCCAACATTTCTCTTAAAGAGGTGGCTGGAGCTAAAGGCATAGTCAAGGTTAATGCT  
CCTTTTTCTTTATCTGACCTCTCCCAATCAGTTAGCGTTTACGCTCTTTTTCATCAAT  
ATAAAAACCCAGCCAGTTCATGGCCATCTGGCAACAACCTTACAGGCTTTACAGCCCT  
AGACCCTGAAGGGTCAAGAGCCGTCTTATTCTCAATATGCATTTTATACCCAATCCGC  
TCCCAACATTAAATAAAGCTCCAAAAATTAAATTCTGGCCCTCAAACCCCAACAGGAC  
TTAATTAACCTCACTTCAAGGTGTACAAGAATAGAGTAGAGGCAGCCAAGTAGCAACGTA  
TTTGAGTTGCAATTCCTTGCCCTCAACTCTGAGAGAAAACCCAGCCACATCTCCAGCAAAC  
AAGAACTTCAAAACACCTGAACTGCAGCAGCCAGGCGTTCTCCAGGACCACCTCCCCCA  
GGATCTTGCTTCAAGTGCCGGAATCTGACCATTTGGGCCAAGGAATGCCTGCAGCCCAGG  
ATTCCTCCTAAGCCACGTCCCATTTGTGACAGGACCCCACTGGAATCGGACTGTCCAAT  
CACC CGCAGCCAATCCCAGAGCCCTGGAACCTGCCCCAAGGCTCTCTGACTGACTCC  
TTCCAGATCTTCTCGGCTTAGCAGCTGAAGACTGACACTGCCGATCACTTCAGAAGTC  
CCCTGGACCATCACGGATACTGAGCTTCAGGTAACCTCTCACAGTGGAGGCTAAGTCCATC  
CCCTGTTTAAATCGATACAGGGGCTACCCACTCCACATCACCTTCTTTTCAAGGGCCTGTT  
TCCCTTTCCCCCATAACTGTTGTGGGTATTGACGGCCAAGCTTCAAAACCCCTTAAACT  
CCCCACTCTGGTGCCAACTTGGACAACATTCTTTTATGCACTCTTTTTCAGTTATCCTC  
ACCTGCCCAGTTCCTTATTAGGCCGAGACATTTTAAACCAATTATCTGCTTCCCCGACT  
ATTCCTGGGCTACAGCCACATCTCCTTGCCGCCCTTCTTCCCAACCCAAAGCCTCCTTCA  
TATCTTCTCTCATATCCCCCACCTTAACCCACAAGTATGGGACACCTCTACTCCCTCC  
CTGGCAACCGATCACACGCCCATTAATCTCCATTAAACCTAATCACCTTACCCTGCT  
CAATGCCAGTATCCCATACCACAACAGGCTTTAAAGGGATTGAAGCCTGTTATCACTTGC

FIGURE 22.1

52/64

CTGCTACAGCACGGGCTTCTAAAACCTATAAACTCTCCATACAATTCCCCATTTTACCT  
 GTCTAAAAACCAGATAAGTCTTACAGGTTAGTTCAGAATCTGCACCTTATCAACCAAATT  
 GTTTTGCCTTATCCACCCTGTAGCACCCAACTCGTACACTCTTTTGTCTCAATGCCTTCC  
 CCCACAACCTCACTATTCCGTTCTTGATCTTAAAGATGCTTTTTTCACTATTCCCCTGCAC  
 CCCTCATCCCAGCCTCTCTTTGCTTTTACCTGGACTGACCCTGACACCCATCAGTCCCAG  
 CAGCTTACCTGGGCTGTACTGCCGCAAGGCTTCAGGGACAGCCCTCATTACTTCAGCCAA  
 GCTCTTTCTCATGATTTACTTTCTTTCCACCTCTCTGCTTCTCACCTTATTCAATATATT  
 GATGACCTTCTACTTTGTAGCCCTCCTTTAAATCTTCTCAACAAGACACCCCTCCTGCTC  
 CTTCAACATTTGTTCTCCAAAGGATATCGGGTATCCCCCTCCAAAGCTCAAATTTCTTCT  
 CCATCTGTTACATACCTCGGCATAATTCTTCATGAAAACACATGTGCTCTCCCTGCCAAT  
 TGGCTCTCCAACCTGATCTCTCAAATCCCAACCTCTTCTACAAAACAACAACCTCCTTTCCC  
 TCCTAGGCATGGTTGGATACTTTTGCTTTGGATACTGGTTTGGCCATCCTAACAATAAT  
 CATTATATAAACTCACAAAAGGAAACCTAGCTGACCCCATAGATTCTAAATCCTTTCCCC  
 ACTCCTCTTTCCATTCTTGAAGACAGCTTTAGAGACTGCTCCCACTAGCTCTCCCTG  
 TCTCATCCCAACCTTTTCATTACACACAGCCGAAGTGCAGGGCTGTGCAGTCGGAATTC  
 TTACACAAGGACCGGGACCATGCCCTGTAGCCTTTTGTCCAAACAACCTTGACCTTACTG  
 TTTTAGGCTCGCCATCATGTCTCCATGCGGTAGCTTCCGCTGCCCTAATACTTTTAGAGG  
 CCCTCAAAATCACAACTATGCTCAACTCACTCTCTACAGCTCTCACAACCTCCAAAATC  
 TATTTTCTTTCTCACACCTGACGCATATACTTTCTGCTCCCCGGCTCCTTCAGCTGTATT  
 CACTCTTTGTTGAGTCTCCCACAATTACCATTCTTCCTGGCCCAGACTTCAATCTGGCCT  
 CCCACATTATTCTGGATACCACACCTGACCCTGATGATTGTATGTCTCTGATCTACCTGA  
 CATTCACCCCATTTCCCATAATTTCTTCTTTTCTGTTCTCTCATGTTGATCACATTTGGT  
 TTATGACGGCAGTTCCACCAGGCCTGATCGCCACTCACCAGCAAAGGCAGGCTATGCTA  
 TAGAATCTTCCATCCATCATTTGAGGCTACTGCTCTGCCCCCTCCACTACCTCTCAGC  
 AAGCCGAACCTGATTGCCCTTAACCTCGGGCCTTCACTCTTGCAAAGGGACTACACGTCAATA  
 TTTATACTGACTCTAAATATGCCTTCCATATCTTGCAACCACATGCTGTTATATGGGCTG  
 AAAGAGGTTTCTCACTACGCAAGGGTCTCCATCATTAAATGCCTCTTTAATAAAAACTC  
 TTCTCAAGGCTGCTTTACTTCCAAAGGAAGCTGGAGTCACACACTGCAAGGGCCACCAAA  
 AGGCGTCAGATCCCATTACTCTAGGAAATGCTTATGCTGATAAGGTAGCTAAAGAAGCAC  
 CTAGCGTTCCAACTTCTGTCCCTCATGGCCAGTTTTTCTCCTTCCCATCAGTCATTCCCA  
 CCTACTCCCCATTGAAACTTCCGCCTATCAATCTCTTCTCACACAAGGCAAATGGTTCT  
 TAGACCAAGGAAAATATCTCCTTCCAGCCTCAGAGGCCATTCTATTCTGTCATCATTTCT  
 ATAACCTCTTCCATGTAGGTTACAAGCCACTAGTCCACCTCTTAGAACCTCTCATTTCCT  
 TCCATCGTGGAAACATATCCTCAAGGAAATCACTTCTCAGTGTCCATCTGCTATTCTAC  
 TACCCCTCAGGGATTGTTCAAGCCCCCTCCCCCTCCCTACACATCAAGCTCGGGGATTGTC  
 CCCTGCCCAGGACTGGCAAATTGACTTTACTCACATGCCCTGAGTCAGGAACTAAAAATA  
 CCTCTTGGTCTGGGTAGACACTGTCACTGGATGGGTAGAGGCCCTTCCCACAGGGTCTGA  
 GAAGGCCACTGCAGTCATTCTTCCCTTCTGTGACACATAATCCTTGGGTGGCCTTCC  
 CACTCTATACAGTCCAATAACGGAGCAGCCTTATTAGTCAAATCACCTGAGCAGTTTT  
 TCAGGCTCTTGGTATTCAGTGGAACTTCGTACCCCTTACTGTCTCAATCTTCAGGAAA  
 GGTAGAATGGACTAATGGTCTTTTAAAAACACACCCCACTCAGCCTCCAACCTTAA  
 AAAGGAGGATAGAGCCCCAAAACTCGCAACCAAGCTAGTAATTATGCTGAACCCCTTGG  
 GCACTCTCTAATTGGATGTCTTAGGTCCTCCCAAATCTTAGTCCTTTAATATCTGTTTTT  
 CTCCTTCTCTTATTCGGACCTTGTGTCTTCCGTTTAGTTTTTCAATTCAACAAAACCGC  
 ATCCAGGCCATCACCAATCGTTCTATACAATAAATGCTCCTTCAACAACCCCAATAT  
 CGCCCTTACCACAAAATCTCCTTCACTTAATCTCTCCACTCTAGGTTCCCATGCCG  
 CCCATAATCCCTCTCGAAGCAGCCCTGAGAAACATAGCCATTATCTCTCCATACCACCC  
 CCAAAATTTTTGCTGCCCCAACACTTCAACACTATTTTACATTATTTTCTTATTAATAT

FIGURE 22.2

54/64

GATCTCTTGATCCCAGGAGGTCAAGGCTGCAATGAGCTAAGATCAAGCCACTGCATTCCA  
 GCCTGAGTGATAGTGGGAGACCTTGTCTTTAAAAACACACACACACACACACACACG  
 AGGGCCTTTGACCACTCTTGAGTAGAAGACTCGAGAAGAACAAGTAGAAGGCCAGAGAA  
 GAACAAAGTTACTTGAAAGATCTCTTATTAAAGAGAATGTACAGCTATGAAAAA  
 AACACACACACACACACAAACCTCATCTGGAATGAAAAAACATAATGCATTGGTTCT  
 GGTTCCTTAGGCTGTTATGGAACAACCAAAGAACATTATTTTGGTTTCTGAGGTGAGAAC  
 TATTTTATCCCTCAAGCACACTATGCTTATGTTTGAGGGAGAATGAGAAATAGGAAA  
 CTAGGAACAGGCTGAAATGGTCTAATCTTGACCATCTAATCTGCAGTGTCTTATCTCA  
 TTCTAAAAGAGAATGGTTATATTCGCTGTTCTAGCATAAAAAGTAATGATAAAAAATAAA  
 GATCCCGTATTACCAGACAATAATCCCTAGACTGTTTTAATGCTTGGTTGAGTATTGTC  
 TTATGATCTCAGACTTTAAAGATGGTCTCCCCCTATGGTGAAGCTTGTAAATTATGTAG  
 GCATCATTAAATGTCTGTTACTTATCAAAATTTTATCATTGTTAGTTGTATTACTACTTG  
 ACAGTCCAATTTATTTAATTGAAAAGATTGGTTAACATTTTATAGTCAAAGTAATTGTTT  
 CCTGTGTTTTTCTGTTTAGGTTATTGGAGTGATGAGTAAAGAATACATACCAAAGGCC  
ACACGTTTTGGACCCCTAATAGGTGAAATCTACACCAATGACACAGTTCTTAAGAAGGCC  
AACAGGAAATATTTTTGGAGGGTAAGTAAGGGAATTTCTTCAGACCCATTAAATGTTAG  
 GAAAAATGGAGCTAAAGAGCTGGGTGGCTCACCTTCTCATCTGTGCTGAGAAATGC  
 TGGGGCTCACCCATAAGTATCCAGCATCCCATGGACACAGGGAATTCTGAACAAATGTG  
 ATGAAACCGATGAAATGTCTGGCCTGTAGGTGGTTAGTGATGGAGATACGGGCTATATGT  
 GAATCTTGATTTTGTCAATTCATTAGAGCTTTGTAATGAAAGGAAACAGTTTGTGTCTTG  
 CTTTAAGGATAGGTTTCATTGTCATTTCTCCGCAAGGAAGTAGTAATGAGTTACCAAGCCT  
 TAGATTTTACCCCTTTTGTATTTCTTGCTGACTTAACTTTAATTGAATGGAAGAGTTATC  
 ACAAATGAATTATCTTTTGGTTTTTTTTTTTTTGAGATGGAGTCTCACTCTGTCACCAG  
 GCTGGAGTGCAATGGCATGATCTCGGCTCACTGCAACCTCCGCCTCCAGGTTCAAGCAA  
 TTGTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTAAGGTGCGCGCCACCATGCCAGTTA  
 ATTTTTGTATTTTAGTAGAGACGGGGTTCCACTATGTTGGCCATGATGGTCTCGATCTC  
 TGGACCTCGTGATCCGCCCACCTTGGCCTCCCAAAGTGCTGGAATTACAGGCAAGAGCCA  
 CCGCGCCAGCCAGGAATGACAAATGAATTACCTTATAAGTAAATGCCATTAAAGGAAGGA  
 TAGCTGGAAGATGGGTTGAGGGGAATGGAGGACCACAGAAGTAGTCCTATTAAATACAT  
 GTGCATGGTAAAATGATTCCATTTGACAATAGGTTAATTATCTCATAGCATAAGGAAAAT  
 GCTTAACAGTCATATGCAAGATGATAAGCTTTCTATAGCATCCAACCAAAGATCTAGC  
 CAGTACAATTTCTTTGCTATATTAGGGTTAGAAAGGCCCCAGAGGTGAACCAATTAGA  
 TGGAAATCCTGAATAAAACACTGGATTAGCAGTGAACAGAAAAAGTCAGATTGCTTCC  
 TTCTTCCCATAGATGTCTCAGGGATATTTAGTTTCTCAGAAGATAAAGAATTTAGTAAG  
 CGTTTTTTGTGCATACTTACATGAAATGTACATTATTTGAATCTTTAAAAAGAAACAG  
 CTGCATGATAACAAAAATTGTGTATGCTTGCTTTAGCTGGTATTTTGCCTAGAACGAT  
 TATATCGTTCGGACAAGAAGCTATTCTAAGAAACAATATTTTAATCCAGGAAGTTTTT  
 CATTTTTAGAAATTTATCTTACTATTTCCCAAGCAAAAGAGGGTAGTTACAGATTTACTA  
 AGAATCATGTGCTCACAATTTTATTTAATAATTATTCCTCCTTAAATATATTAATCAC  
 CTGACTTACAATGGTGAACCATGAGTGCAATTTTGCCTTTATTGTCAATAACGTCTTCT  
 CAGAAGTGAGCCACAAAGGTGCATAGTTCTTGAGGTTAAAGGTCTGAATTAAGACAATCC  
 AGCATAAGTCTCATTAAATGTGTGATTATTTGAGAAAAGGCAAGAAGTACCTAAGAATCT  
 CCCCTCACTGTCCAGTTCCCTGTTTCATTTAAAGATTCACTGTAAGTAACTGAAAGGCT  
 TTCTTGGGAGGATTTATTTGAATCAGTCTTTCACATGCAAAGGATATTGTAGAACATCT  
 CGTTTTTGTGCTGGCAGGAATATGAACATCTGTTGTGAGGAAAGAAAAAGTTTCATGCAAT  
 TACACTGCCAAAGAAGGGATGTTCAAGTTGAGAAACCAAGTGACATTTCTTGTAAGTGTAC  
 TATGAATCAGCGCATTTTAATCTTCTAGATAATATATGGAAGTGACAGGAAGGTGGTAGGA  
 AACGGTGTTCATTTTACATATGCGTTATTTTATTCTGTGTGAGTGACTTCATGGCACCGA  
 CATTGCTGTTTTTAAATGAGGATACAGTAAATTGCAGTCCGAGGAAGGCTAACTGGAATC  
 AACATACCGTAGCTTTAGAAAGCAGTTTCCGCACCAGCGAAGAGTACAAGAGCGATGGA  
 ACCCATGTTCTGGAAGTTTGCACATCAGAGTAAACAACTTGAACCCCTCTTGATA

FIGURE 23.1

55/64

GCAGAAATTCACCCAGCCTTGTTCCATTTTCTCTTAACAAAACACACCGCAAAAGCTCTCA  
 CAAGCTGCTTTGATGAAGCCACATGTATTTCCCCCTTACAATTTACAGGAAGTTACTCT  
 TAAAAGAAAGTGATTCTGGTGTTCACGCCTGTGTTAAAGGGACAGAGTTCCTTTTTATT  
 TCTGATAACGTTTGAGCGAAATACAGAACTATCTGTAGACTAGCATAGTCGGTACGTGA  
 GTAAGGAAAAGCAATAACCTGCTGTCCGGTGAGCACAAAATTCCTGCTACGAACAGTGCC  
 TTACTGCTGCTTGAGACTGCAAGTCGCAGATCACACTAGGTATTGACTGATTGTATAAG  
 GAAATTTCTTAAAGTCTAAAGTAAAGGTGGTACCTCCTAAAAAGAGGGGAAGAGAGAAAA  
 CTTTGTGTGGAAAGGTTAAGGAGTGTGTTATAGTTTCAGTAAGAGTGTACGTTTAAATTT  
 TTCTTCTTCTCTGCCTCTTTGCCAAGTAGCCTGAGTGCATCTGTTATCCAGAAGTAGTA  
 TTACTCTAGGACAAACTTCAAATTCTTCATTCTGCGTTGCCTTTAAGGAACAACATACTT  
 TCTTCTGTCTTTTTTCCAAAACACACGCCTATGGCTCTGTGTGTGGTGTTTAGCCAG  
 CCTCCTCCCAGATAAGGGGTTCCTTCCCTCCTTTGCATTGAAAGGAAAGTGCAAGCTG  
 GACATGTTTATCAAGAGGAAAAGTGACTTCTCAGTAATAGACTGTCAAATTCGGGCTGCT  
 GCCCAGAGTGTTCGCTTTGTTATGGCAGGTGAAGTTCACCTTTGCCCCACCCAGTGTTC  
 ACAAAAAGGCAAGGTTCCAGTATTCATATGAACAAGTGTACTTTAGGACTTGGAGGGT  
 TGGGGGTGGAGGATGTTTGCATAGTTGAAGCCTTGGGCGGGGTGTAGGAAACGGCGAGT  
 ACAGAGGCCATAGAAAAGCTAAGACTCAGTTTGACGTCGTCAGCCGGCTTGGTCTTCTA  
 CCCAGTGACTCAAAGCACTAAAAGTCAGCATAATCGGAACTGAAGTCAGTAGCATCGCCC  
 ATTTGCCATTCACTGCAGTAGCAAAAGTAGTACTCTGTGGTGGGTAAATCGGTTTGAGGC  
 AGCTCCTTAAATGAACATTTGTGTTTCATTTTCTGTTATTTTCCCGAACATGAAAAGAC  
 GATAAACTGAAATGGAAAAGGTAAGTGCACAAAAGTGCCTTACCTGTTTCCGCCCTGA  
 TTTCTGCTGATTCAGACTATTCTGGCTAACTGATTGGATTCTTTTCTAACTAGGCAG  
 TAGGGGATCAGAAATCACACACGTAACGGCTGTGTTTATTCTGAGAGGTGCTGGGGAGC  
 TTTGGGTCTGACTTCTTTTACATGCCTGTCTTCTTTTGGACAGATCTATTCCAGAGG  
GGAGCTTCACCACTTCATTGACGGCTTTAATGAAGAGAAAAGCAACTGGATGCGCTATGT  
GAATCCAGCACACTCTCCCCGGGAGCAAACTGGCTGCGTGTGAGAACGGGATGAACAT  
CTACTTCTACACCATTAAGCCCCATCCCTGCCAACCCAGGAACCTCTTGTGTGGTATTGTCTG  
GGACTTTGCAGAAAGGCTTCACTACCTTATCCCGGAGAGCTGACAATGATGAATCTCAG  
 TAAGTGGATTACAGAACAAAAATAAAAAATGCCAGTAATGTGGTTCTGCCCTTTGA  
 ACTAATAACATGTTGTTTAAATTATACGGCTTTGTCTGTGTTGGATGAAGTAGGTGGCTT  
 AAGCTAGGGACTAGGAAGAGGAAAAACATTTTTTGTAGTCCCTATTAATATTAGGAACT  
 TGATCATTTAAAGTATATATATATATAGAGAGCTACCTTGAGTTTGAATTCAGGATGT  
 TACAGGAAGAAATATATGTCCAATTCTAATTTATCCAAAAGCAGTTGGGAGAAATACAGG  
 GATTGGTCCAGACATGCTGCGTATGCAAGGTATAGCCCTCATCTGTGGTACTTTGGCAGG  
 GCTTAGACTGCATCAAAATATTTATAGATGTACATTTGAGTGTACAGTTAGGATCTGATG  
 TGGAACATTGTAAAGATCATTGCTAGAAAACTTTGTCTAATTTTCAATATTATCTAA  
 GTGAATAACCGTAAAGATTTTACATCTTAGCTTCTTCTTACAGTAAAAAACTATCTG  
 ATCTCTTGATCAGTATTATAGTAGCCACCTATCACTTTATCTTAACAAATTCTCAATTCC  
 TTAGGTTTATGTGCTTTTACTTCTTTTATTGATTAAAATTGCTGTCTGACCTCTCTCT  
 GCAGAGGGCTGCATCATTTTGGTCATTCTCAAGTGATCTCTTTGAGCAATTTAAGAATTG  
 CCATAAGATTCTAACCTCTGCTGTAACATATGGTTGTGTGTTCTTGGTTAGACCACTAAAT  
 CTTATTAGCAGTTTTAAAAATTATTCCTTTTGGTTTAGAAGTTAAGACTAAATGCTGAAG  
 TTTTGTAACTTTTGGTTTGTATATCATTTCAAACCTAAGAAAACATTTGAAGAAAAGGA  
 CAAAGAATTTCCACTTACCTTTTACCCAGGTTTACCAGTTATTGATAAGTATATCCATTT  
 GCTTTACCAGAAGGCTAACTTGTGTTTGTCTCATTTTACCTTTGAGACATTTGGAATA  
 AATATCAATGTTAACATAAATTGGAATTTTGACTTTGATTTTAGGACCAATGAACAAGCC  
 AAGTACTTACCTAGTCATATATAATCCAAGTGTATGGTTATTGTTGATTCAATCCACAC  
 TTCATTTTACTTGTATCTCCCTTAAGATTGCAAGATTGTGTTTGCAGTTTTTCTGAAAATC  
 TGGGGCTATAAAAGCATCAGGACCTCCCCGTAGGGGAGGTGCTGTGTTTGGGGTCTTA  
 CACAACAGGTTACCTTGAGCTTCAGGAAAAGAACTGGCTCTCAGTTCCCAAGTCCAGC  
 TTAATGGGTCTAATTAGGTCTGACCAAAAGGTGGCAGTTCTTTCCCTCATGTCTCTT  
 CAGCGCTCCCCGAGACTCTGGAGACTCTGTCTATATCCCTAGGGCTGAGCCTCCAGGAAC  
 CATTGGGCTGTTGTGGCATCTGTGTATGCCATGCCAGTGTGAGGACCTAGTAACAAAC

FIGURE 23.2



56/64

GACAAATGCACAGGCACAGTGGCATT TTTGTGGAAC TCGTATTCCAGCTGTGCGTCTCAG  
AAGAAGCGCACAGCTCCCTCCTGGCTTTCTTAACATAGTGAGCCACTTCCACTTAAGGGT  
CTCCTTACATTCCCTTGAGTTTAATCATT CATGGATT CAGAGGAAAGTCTTTTGATTTTG  
CTTTCTTTAAACAGTTCATTGAGGTGACCTACCCAGTGACTTTGCACCAACCACCAA  
GAACTTTTTTG CATGCTTCCCGCACCCCTGTGCCAATCAAGGGAAGGGTTTAAAGGCCTG  
GCGTTTTTATTCTCTCAAAGAAAGGTTTGCACAGTATTTAAGGTTCAAGTGCTTCTACT  
TTGTGTT CAGAAGCAACTGTCATATATACTGTGAAATGACACCTTTTATTATCCCTTTT  
TATTTATGCAGTATGTCCCTTTTATTTTGGCAGAA TTTTCTAAATGGTGGTTTAAACA  
TTTTCAAGCACATTTCATTGTCCAATATT CATAGTAAAGAATGAGAGTTAACAATAACCA  
GTCACATTAAAACAAGATTCTGTGCTGCCAGTTGTGAAACCGGTTGTCTTAGGCGTGGCAG  
CTGATGATTGAGACTGTGATCAGGAAAATTTCCACTATTT CATCAGGCCTAATAGGTAGA  
TTGTGTCTCCAAATGAAC TGTGTTGGGTTTCCATGCTTAAAGCACAATAGAGGTGGTGCA  
AGAATCTCCATGAGGGCTTAAATGGCAGTGTGTT CAGGCGGTAGAGTTTGGAGAAGAA  
GGGATTTGAAACCAAACCAAAGGAAAGAAAGTAAGTAGCCAGAAATCACAAAATGGCATT  
TTTCTAAAAACAAAGGAAAAGGAATAAAGAACTAATAAGTTTGAACCCCTACCCCTCC  
CAAAATTTGGCAGGGGGGGAGGTATTTTCTATCTATCTA ACTAACCCATCTAGAAAA  
CAGTTGACCAAAATTATAGACTTCTAAATGTTAATCTGCTTCTCAGTTTCAGTTGAAAAG  
AGACTTTGTTTTGCTACTGCAGAACTTCTAGGTTCTTTCTTATAGTCTTGGGGTTCTTA  
TTATAGATCGAAAATGTGAGTCGGCATAATTAAGCCATT CGGAGTCTTCAGAAGCAGTTC  
ACTCTTGAATGACTCCGTCGCCCTACAGCCATTTAAGATTTCAGAACAAAACAGATCT  
TGATTTTCTTTTCTATGTTAACTCAAGCTGTTGCTGAGTGGGAGAGTCAGAAATGACACC  
AGCTCCACTGATTACTCAGCTGCTGAAGGATGATTTT TAAAAATGCACCTTTACTGTATA  
TGGACTTCTTAATTTCCACCTGTAGAGCATCTTAGGGAGGCTAACATGTCACTCTGGATG  
TTCTTTTAGAATAAGATGCAAATCTATTTTCTGAAGGCATTAGAGATAGCAAACATTTA  
TTGTGAGTTTACTATATACTAGGCACTGTGCTAAGTGT TTTGCATAGAAAGTTTAAAT  
CTGGCTTTTTTGTGGCCCAATCATAAGTTT CATATCAGTTCAACATTCAAATTATATTA  
AGGTACTTAAGAAGAAATCCCTGGCTAAATGTGAGGGGCAGTCCACAGATGGACTGAAAC  
TTTATGCTTATTGCACATTTATGCTATTATTATTTGTTGAATTATAGAACCAAGGGAGTG  
TGGAAAGCCACTGGA AAAAATATGAGACTTAGATACATAATTTGAGTAAAAATGGCTCAAA  
GTCATGAGGGTAAAGTTTTTTGTATTTCCATTTTATTCGAGCGGCATCGTTTTTAAAAAT  
CATTATGAATTTGACCCTATATAGATGTTTCAAATAATTCTTTTCCCTTCATAAAAT  
TCCTTCCTGTGGCTGTGAGATGCCCTGCCTATCAGTTTTCAGCTTAGTTGTCTTTCTCA  
TCCTTTACCAITTTAGCTTTAAAAAACAAAAGTGACAATTAGAACTTCCTGCCTGTGGG  
CCTCACTGAAAGACCGATATTGGCCTGATAAGGAGATATTTATTTGTTTTAGTGGCTTC  
AGAAATCCCTCTCCCTCAGCAAGCTTCCATCACGGCCCCCGTCAGCATCTTCCTGA  
TAGCGTTCTTCTCTGTGTTTATTCTGGGGCTTCAGGCTCGCCCAGGAGGAAGTATAACC  
GCTGGCAGGAGATAACATTCTCTAAGGGGCTCTCAAATTGGAATCGAATCCCTCAAGCCA  
GTCAGCCTAGAGAATACATTTAAAGGGTTCAGTTCTGGAGTTTCACAGAGTTCAATTCTA  
GACCTATCAGATAGCAAGTGTGGAGTTCTTTCTCAACTAAATTCAAGCAGAGACATTTT  
TAGACGATGAAGGATATTTGCACAAAGGCTTCAGCATGATCCCCAAACCTGCTGCCTCT  
GAAGGCATCTCCACACATTGACAGCCAATGCCTTCAGTGCGTTCCTAGGGCAGGTGTCCT  
GGCTTGAGTGACTGTCTCCAATAATCAGAGCTCAAAC TAAACATCGTATGTTTTACTTT  
TGGTTTCCAGGCAAGGCTGAGCAGGGAATTTT CAGTTTTCCCTGCCAGATGGGTGTTTT  
TTCCTGAAGGCATCATTTATTGTGTAGCGAGGAGACAGGGCTGGCTGTGGCAGGGATAGT  
CTAGAAGTGTCTCATTGCTGTGTTCTTAAATAGTATCTTTACCAAGTAATAACGTGCC  
GTCTTTGGGAATAAGTGCTTTCTCTTAGCCTGTCTGTTTTCTTGGGTGCGCTAAGTAA  
TTGAACTGGCTCAGGAAGTACCTATTGTGTTTGGCAGAGGTGACTGTCACGCCTTGTA  
CTCCAGGGGCCAGCACTGTCTGGGATCCTGGCTAGACCAGACAGAGCCTTGGTGAAGTGCT  
TAGGCTGTCTGCACATCGCGAGGAAGGTGGTATTCACTTCGCTAAGCTCCTTGGCATAGG  
CAGTTTGAACAGGGCTTTATCAAATTCGTATTCAACAAGAGTAGAAGCGAAAATTGATGA  
CTGTGTATTACTTGAAATGAGTCTTAATCTTTCACATTTAGTTCTCAGGGTATGCTGATT  
TCCTTTAGGTAACCATGAACATCAGAAAGACTTTTATTAACCTATGACAGGGTCCCCAC

FIGURE 23.3

57/64

CCCAGTATTTTCCACTCCATTAAAATGGAAGTTTTTTTTTTTTTTCTTTTTGAGAC  
AGAGTTTTGCTCTTGTTGCCAGTCTGGAGTGCAATGGCACAATCTCGGCTCACCACAAC  
CTCCACCTCCAGATTCAAGCGATTCTTCTGCCTCAGCCTCCCAAGTAGCTGGGATTACA  
GGTGTGCCGCCACCACGCCAGCTAATTTTGTATTTTAGTAGAGATGGGTTTCTCCATG  
TTGGTCAGGCTGGTCTCGAACTTCCGACCTCAGGTGATCCGCCACCTCGGCCTCCCAA  
GTGCTGGGATTACAGGCAAGAGCCACTGCATCCAGCTTAGGCTATCTTACTCCAGCCTAA  
ACAGCAATTTTCTATCATAGGTCTGTACTAATGAAAACAGAATCACCCAAGGCTGCTGT  
TTGTTCTGTCTGTGCTGCCATTGTCCGATTTTGCTGAGGAGGAAACGGAATGCATTT  
TGAGTGAGTGGCCAGAGCCTTCTAGAATGAGAGTGCCTTGGAAAGCCAGATATGTGGCGA  
TTGTGTGCCAGCTGTTACTCAGGTTTTCTCAAGAAGGAGGAGCAACTTTGGCAGTTTTG  
CTTCAGTTCTCTAGCCCTCTGTGTAATCGCCCTTTTTCTTTATTTAGCACAAACAC  
AGAGCAGTCTAAAGCAACCGAGCACTGAGAAAAATGAACCTCTGCCCAAAGAATGTCCAA  
AGAGAGAGTACAGCGTGAAAGAAATCCTAAAATTGGACTCCAACCCCTCCAAAGGAAAGG  
ACCTCTACCGTTCTAACATTTTCAACCCCTCACATCAGAAAAGGACCTCGATGACTTTAGAA  
GACGTGGGAGCCCCGAAATGCCCTTCTACCCCTCGGGTCGTTTACCCCATCCGGGGCCCTC  
TGCCAGAAGACTTTTTGAAAGCTTCCCTGGCCTACGGGATCGAGAGACCCACGTACATCA  
CTCGCTCCCCATTCCATCCTCCACCACTCCAAGCCCTCTGCAAGAAGCAGCCCGACC  
AAAGCCTCAAGAGCTCCAGCCCTCACAGCAGCCCTGGGAATACGGTGTCCCTGTGGGCC  
CCGGCTCTCAAGAGCACCGGACTCCTACGCTTACTTGAACCGCTCTACGGCACGGAAG  
GTTTGGGCTCCTACCTGGCTACGCACCCCTGCCCCACCTCCCGCAGCTTTCATCCCT  
CGTACAACGCTCACTACCCCAAGTTCTCTTGCCCCCTACGGCATGAATTGTAATGGCC  
TGAGCGCTGTGAGCAGCATGAATGGCATCAACAACCTTGGCCTCTTCCCGAGGCTGTGCC  
CTGTCTACAGCAATCTCCTCGGTGGGGGAGCCTGCCCCACCCATGCTCAACCCCACTT  
CTCTCCCGAGCTCGCTGCCCTCAGATGGAGCCCGAGGTTGCTCCAGCCGGAGCATCCCA  
GGGAGGTGCTGTGCCCGGCCCCACAGTGCCTTCTCCTTTACCGGGCCGCCCGCAGCA  
TGAAGGACAAGGCCTGTAGCCCCACAAGCGGGTCTCCACGGCGGGAACAGCCGCCACGG  
CAGAACATGTGGTGCAGCCCAAAGCTACCTCAGCAGCGATGGCAGCCCCCAGCAGCGAGC  
AAGCCATGAATCTCATTA AAAACAAAAGAAACATGACCGGCTACAAGACCCCTTCCCTACC  
CGCTGAAGAGCAGAACGGCAAGATCAAGTACGAATGCAACGTTTGGCCCAAGACTTTCCG  
GCCAGCTCTCCAATCTGAAGGTAGGCCTTGAGAGAGAGCAGTCCAAGGGGCTGTGAGTGC  
ATGCTTGTGTTTGTATTTAGCTTGCTTTCCATGGGGTATCGATTGCATTTGCAGTAGTAT  
GAGCCCCCGTTGGGGATAGTGGGTATGGATTCCGCCTGGCTTTTGCCACTTCTAGCTCT  
TTGACTTTGGACAAGTGACTTCCCTTCTCCTGATTTTCTTGAATAATAAAAAAATTAG  
GGGTTTGGACTAGAAGATTAGGTGAAACTCCCTGCTAGCCTGTGATTTTGTGCTTTTAA  
GAAAAACACCATTTCTGAAAACATGAAGATTCTTCTTTTAAAGACTGTCTTGATGCTTTT  
CTTAAGATATTTGCATCAACACTTGAGTCTTGGAGCAGAAATGTTAGGTCTCAGAGCCAG  
CTTGAGAGCAGAGCTAACACATGTGGCTTCTCCAGGTCCACCTGAGAGTGACAGTGG  
AGAACGGCCTTTCAATGTGAGACTTGCAACAAGGGCTTTACTCAGCTCGCCACCTGCA  
GAAACACTACCTGGTACACACGGGAGAAAAGCCACATGAATGCCAGGTGCGCAGTATTTT  
CTGGGTAGACCTTCTGACCTTTGTAGAAAATGTCTGTGAGTCACCTCCCATGTCTCTATA  
TAGCCCGTAGTTAAAGCCAACACCAGATTCTGCGTTGTCCCATCCTGGACTGATGGCACT  
ATGGTCTTCCAGTACTTTGTATCTGCTGATGACTTGAGATGGCACAGCCAGCTTCCAG  
TGGGTGGGAAAATGGTAGGGGAAATAAACAGCCCTCGTGTGCTGTGCCACATCCCC  
CCGTTTGCTTAATACCACACTGGAGGTGCCACAAGGAGGCTTCTCACCTCCTAGGTTGCT  
GGGCGTTGGCGGTAAGCCTGCCCCTCCCGTTGGCAACTCTTAATCTTCTGGCCTTCTG  
TCTCCCTTCCCTGCTGTCTCTCTCCCTACACTGTAGGTCTGCCACAAGAGATTAGCAG  
CACCAGCAATCTCAAGACCCACCTGGCAGTCCATTCTGGAGAGAAACCATAACCAATGCAA  
GGTGTGCCCTGCCAAGTTTCAACCCAGTTTGTGCACCTGAAACTGCACAAGCGTCTGCACAC  
CCGGAGCGGCCCCACAAGTGCTCCAGTGCCACAAGAACTACATCCATCTCTGTAGCCT  
CAAGGTTACCTGAAAGGGAACTGCGCTGCGGCCCGCGCTGGGCTGCCCTTGAAGA  
TCTGACCCGAATCAATGAAGAAATCGAGAAGTTTGACATCAGTGACAATGCTGACCGGCT  
CGAGGACGTGGAGGATGACATCAGTGTGATCTCTGTAGTGGAGAAGGAAATCTGGCCGT

FIGURE 23.4

58/64

GGTCAGAAAAGAGAAAGAAGAACTGGCCTGAAAGTGTCTTTGCAAAGAAACATGGGGAA  
TGGACTCCTCTCCTCAGGGTGCAGCCTTTATGAGTCATCAGATCTACCCCTCATGAAGTT  
GCCTCCCAGCAACCCACTACCTCTGGTACCTGTAAAGGTCAAACAAGAAACAGTTGAACC  
AATGGATCCTTAAGATTTTCAGAAAAACCTTATTTTGTCTTAAAGTTATGACTTGGTGA  
GTCAGGGTGCCTGTAGGAAGTGGCTTGACATAATCCCAGCTCTGCAAAGCTCTCTCGAC  
AGCAAATGGTTCCCTCACCTCTGGAATTAAAGAAGGAACCTCAAAGTTACTGAAATCT  
CAGGGCATGAACAAGGCAAAGGCCATATATATATATATATATATCTGTATACATATTA  
TATATACTTATTTACACCTGTGTCTATATATTTGCCCTGTGTATTTGAATATTTGTGT  
GGACATGTTTGCATAGCCTTCCATTACTAAGACTATTACCTAGTCATAATTATTTTTC  
AATGATAATCCTTCATAATTTATATACATTTATCATTGAGAAAGCAATAATTAAAAAA  
GTTTACAATGACTGGAAGATTCTTGTAAATTTGAGTATAAATGTATTTTGTCTTGTGG  
CCATTCTTTGTAGATAATTTCTGCACATCTGTATAAGTACCTAAGATTTAGTTAAACAA  
TATATGACTTCAGTCAACCTCTCTCTAATAATGGTTTGAAAAAGAGGTTTGGGTAAT  
GCCAATGTTGGACAGTTGATGTGTTTCTTCTGGGATCCTATCATTGAAACAGCATTGTA  
CATAACTTGGGGGTATGTGTGCAGGATTACCAAGAATAACTTAAGTAGAAGAAACAAGA  
AAGGGAATCTGTATATTTTGTGTAGTTCATGTTTTTCCCCAGCCACAATTTTACC  
GGAAGGGTGACAGGAAGGCTTTACCAACCTGTCTCTCCCTCCAAAAGAGCAGAAATCCTCC  
CACCGCCCTGCCCTCCCCACCGAGTCTGTGGCCATTGAGAGCGGCCACATGACTTTTGC  
ATCCATTGTATTATCAGAAAATGTGAAGAAGAAAAAATGCCATGTTTTAAACCCTGC  
GAAAATTTCCCAAAGCATAGGTGGCTTTGTGTGTGTGCGATTTGGGGGCTTGAGTCTGG  
GTGGTGTTTTGTGTGTGGTTTTTGTGCTTTTTTTTTTTTTTTTTTTTAAATGTCAAAAT  
TGCACAAACATGGTGCTCTACCAGGAAGGATTCGAGGTAGATAGGCTCAGGCCACACTTT  
AAAAACAAACACACAAACAACAAAAACGGGTATTCTAGTCATCTTGGGGTAAAAGCGGG  
TAATGAACATTCTATCCCCAACACATCAATTGTATTTTTCTGTAAAACCTCAGATTTTC  
CTCAGTATTTGTGTTTTTACATTTTATGGTTAATTTAATGGAAGATGAAAGGGCATTGCA  
AAGTTGTTCAACAACAGTTACCTCATTGAGTGTGTCAGTAGTGCAGGAAATGATGTCTT  
ATCTAATGATTTGCTTCTCTAGAGGAGAAACCGAGTAAATGTGCTCCAGCAAGATAGACT  
TTGTGTTATTCTATCTTTTATTCTGCTAAGCCCAAAGATTACATGTTGGTGTTCAAAGTG  
TAGCAAAAAATGATGTATATTTATAAATCTATTTATACCACTATATCATATGTATATATA  
TTTATAACCACTTAAATTGTGAGCCAAGCCATGTAAAAGATCTACTTTTTCTAAGGGCAA  
AAAAAAAAAAAAAAAAAAGAACCTCTTTCTGAGACTTTGCTTAATACCTGGTGACC  
TCACAATCACGTCGGTATGATTGGGCACCCCTTGCTACTGTAAGAGACCCTAAACCTTG  
GTGCAGTGGTGGGGACCACAAAACAACCAGGGAGGAAGAGATACATCATTTTTTAGTATT  
AAGGACCATCTAAGACAGCTCTATTTTTTTTTTGGCCACTTATGATTATGTGGTCACACC  
CAAGTCACAGAAATAAAAACTGACTTTACCGCTGCAATTTTCTGTTTTCTCCTTACT  
AAATACTGATACATTACTCCAATCTATTTATAATTATATTGACATTTTGTTTACATCA  
ACTAATGTTACCTGTAGAAGAGAACAAATTTGGAATAATCCAGGGAAACCAAGAGCCT  
TACTGGTCTTCTGTAACCTCCAAGACTGACAGCTTTTTATGTATCAGTGTTGATAAACA  
CAGTCCTTAACCTGAAGGTAAACCAAGCATCACGTTGACATTAGACCAATACTTTTGAT  
TCCCAACTACTCGTTTGTCTTTTCTCCTTTTGTGCTTTCCCATAGTGAGAATTTTAT  
AAAGACTTCTTGTCTCTCTCACCATCCATCTTCTTTTCTGCTTTACATGTGAATG  
TTGAGCCCAACATCAACAGTGGTTTTTTTTTCTCTACTCAAAGTTAAACTGACCAA

FIGURE 23.5

**FIGURE 24**

60/64

GAATTCGGGAAGCCAGACGGTTAAACACAGACAAAGTGCTGCCGTGACACTCGGCCCTCCAGTGTGCGG  
AGAGGCAAGAGCAGCGACCGCGCACCTGTCCGCCCGGAGCTGGGACGCGCGCCCGGGCGCGGACGAAG  
CGAGGAGGGACCGCCGAGGCTGCCCCAAGTGTAATCCAGCACTGTGAGGTTTCAGGGATTGGCAGAGG  
GGACCAAGGGGACATGAAAATGGACATGGAGGATGCGGATATGACTCTGTGGACAGAGGCTGAGTTGAA  
GAGAAGTGTACATACATTGTGAACGACCACCCCTGGGATTCTGGTGCTGATGGCGGTACTTCGGTTCAAG  
CGGAGGCATCCTTACCAAGGAATCTGCTTTTCAAGTATGCCACCAACAGTGAAGAGGTTATTGGAGTGAT  
GAGTAAAGAATACATAACCAAGGGCACACGTTTTTGGAGGATCTATTCCAGAGGGGAGCTTCACCACTTCATTG  
ACGGCTTTAATGAAGAGAAAAGCAACTGGATGCGCTATGTGAATCCAGCACACTCTCCCGGGAGCAAAA  
CCTGGCTGCGTGTGAGAAGGGATGAACATCTACTTCTACACCATTAAAGCCCATCCCTGCCAACCAGGAA  
CTTCTTGTGTGGTATTGTGCGGACTTTGCAGAAAGGCTTCACTACCCCTATCCCGGAGAGCTGACAATGA  
TGAATCTCACACAAACACAGAGCAGTCTAAAGCAACCGAGCACTGAGAAAAATGAACTCTGCCCAAAGAA  
TGTCCCAAAGAGAGAGTACAGCGTGAAAGAAATCCTAAAATTGGACTCCAACCCCTCCAAAGGAAAGGAC  
CTCTACCGTTCTAACATTTTACCCCTCACATCAGAAAAGGACCTCGATGACTTTAGAAGACGTGGGAGCC  
CCGAAATGCCCTTTTACCCCTCGGGTCGTTTACCCCATCCGGGCCCTCTGCCAGAAGACTTTTTGAAAGC  
TTCCCTGGCCTACGGGATCGAGAGACCCACGTACATCACTCGCTCCCCATTCCTCCCTCCACCACTCCA  
AGCCCCCTCTGCAAGAAGCAGCCCCGACCAAGCCTCAAGAGCTCCAGCCCTCACAGCAGCCCTGGGAATA  
CGGTGTCCCTGTGGGCCCGGCTCTCAAGAGCACCGGACTCTACGCTTACTTGAACGCGTCTTACGG  
CACGGAAGGTTTGGGCTCCTTACCCCTGCTACGCAACCCCTGCCCCACCTCCCGCCAGCTTTTATCCCTCG  
TACAACGCTCACTACCCCAAGTTCCTCTTGGCCCCCTACGGCATGAATTGTAATGGCCTGAGCGCTGTGA  
GCAGCATGAATGGCATCAACAACCTTGGCCTCTTCCCGAGGCTGTGCCCTGTCTACAGCAATCTCCTCGG  
TGGGGGAGCCTGCCCCACCCCATGCTCAACCCCACTTCTCTCCGAGCTCGCTGCCCTCAGATGGAGCC  
CGGAGGTGTCTCCAGCCGGAGCATCCAGGGAGGTGCTGTGCCCGGCGCCCCACAGTGCCCTTCTCCTTTA  
CCGGGGCCCGCCAGCATGAAGGACAAGGCCTGTAGCCCCACAAGCGGTCTCCACGGCGGGAACAGC  
CGCCACGGCAGAACATGTGGTGACAGCCCAAAGCTACCTCAGCAGCGATGGCAGCCCCCAGCAGCGAGAA  
GCCATGAATCTCATTAATAAACAAGAAACATGACCGGCTACAAGACCCTTCCCTACCCGCTGAAGAAGC  
AGAACGGCAAGATCAAGTACGAATGCAACGTTTGGCGCAAGACTTTCGGCCAGCTCTCCAATCTGAAGGT  
CCACCTGAGAGTGACAGTGGAGAACGGCCTTTCAAATGTCAGACTTGCAACAAGGGCTTTACTCAGCTC  
GCCACCTGCAGAAACACTACCTGGTACACACGGGAGAAAAGCCACATGAATGCCAGGTCTGCCACAAGA  
GATTTAGCAGCACCAGCAATCTCAAGACCCACCTGCGACTCCATCTCTGGAGAGAAACCATACCAATGCAA  
GGTGTGCCCTGCCAAGTTTACCCAGTTTGTGCACCTGAAACTGCACAAGCGTCTGCACACCCGGGAGCGG  
CCCCACAAGTGCTCCAGTGCCACAAGAACTACATCCATCTCTGTAGCCTCAAGGTTACCTGAAAGGGA  
ACTGCGCTGCGGCCCCGGCGCCTGGGCTGCCCTTGGAAAGATCTGACCCGAATCAATGAAGAAATCGAGAA  
GTTTGACATCAGTGACAATGCTGACCGGCTCGAGGACGTGGAGGATGACATCAGTGTGATCTCTGTAGTG  
GAGAAGGAAATTTGGCCGTGGTCAGAAAAGAGAAAGAAAGAACTGGCCTGAAAGTGTCTTTGCAAAGAA  
ACATGGGGAATGGAATCCTCTCCTCAGGGTGACGCTTTATGAGTCATCAGATCTACCCCTCATGAAGTT  
GCCTCCAGCAACCCCACTACCTCTGTACCTGTAAAGGTCAAACAAGAAACAGTTGAACCAATGGATCCT  
TAAGATTTTCAGAAAACACTTATTT

FIGURE 25

**FIGURE 26.1**

62/64

GCTACTGCCACCGCCACGGCCACCACCACACTACTACCCTATCCACCATCACCTCTACCATCA  
CTACTGGCCTCATGGATAGCAGTCACTGGAGATGACGTCTGGGCGGCTCTGCCCTTCTATCCAGCAG  
CAGCACTAATGTCCGGAGACCCAAAGCTCACTTTTGATGACTCGGTTACAATGCTGATTATTACATGCAA  
GAAGCTAAGAAGCTGAAGCACAAAGCTGATGCACTGTTTCGAGAAATTTGGCAAAGCTGTGAATTATGCTG  
ATGCCGCCCTCTCTTCACTGAATGTGGCAATGCCATGGAACGCGACCCTCTGGAAGCAAAGTCCCCATA  
CACCATGTACTCTGAGACTGTGGAGCTCCTCAGGTATGCAATGAGGCTGAAGAACTTTGCAAGTCCCTTG  
GCTTCGGATGGGGACAAAAGCTAGCAGTACTATGCTACCGATGTTTATCACTCCTCTATTGAGAAATGT  
TTAAGCTGAAGAAGGACCATGCTATGAAGTACTCCAGATCACTGATGGAATATTTTAAGCAAATGCTTC  
AAAAGTCGCACAGATACCCTCTCCATGGGTAAGCAATGGAAGAACAACCTCCATCCCCAGTGTCTCTCAAC  
AACGTCTCCCCATCAACGCAATGGGGAACTGTAACAATGGCCAGTCACCATTCCCCAGCGCATTCAAC  
ACATGGCTGCCAGCCACGTCAACATCACTAGCAATGTGTTACGGGGCTATGAACACTGGGATATGGCCGA  
CAAACTGACAAGAGAAAAAAGAAATCTTTGGTGATCTGGACACGCTGATGGGGCTCTGACCCAGCAC  
AGCAGCATGACCAATCTTGTCGCTACGTTGCCCAAGGACTGTGTTGGCTGCGCATCGATGCCCACTTGT  
TGTAGTGGGTGTTCTCAGATCTCTAGCATCACGACCCATCACTCTACCTCTACCAGCGCACTGATGTC  
CTGGTGGAACTCCACTCACTGGGGAACGTTCTCTTTGGTTATGTTTGTGTTTATGCTTCTTTGTTATCT  
GTAAGAAACAGAAGTCATTGTAAGTTGACACTACAACCTAAGGGCAGTGACGTTTATTACTTAGTCAT  
TTTTTTCTTTTAGCATTTGATATGCAATTTCTCAGATTCCACCATCTTTTGTGCTTTATGGAATGACAG  
TCCCTACAATATTGTTTTAAGCCCACTACCCAAAACAAAGAATGGGAAGCACTTGTGATAAAGACAGG  
CTCCTGAGAAATGCAACAAGTGGTCTTACATATACATGAGAACTTAGACACAAGGGACCATCCCCAAAC  
TCTACTCTTATACCCAGAAAAAGAACATATTTTCAAGATCTGTCAAACCTTTGTGTATCCACAGATTTC  
CAATCTTCAGGTGAGAATTTTCAATGTCAAACCCACTGGTTAGATGTTGTAGCAACATCATAAAATCAAGAGT  
ATCAAGAAAATAAATGAGCATAGCAATGCTACTCTTAAAAAGATGCTATGCCACACAACAGAGGACTTT  
CTGTGTAGCATCCCTTTCTGATTCCCTATTTTGTAAATTTAATGATAAGAAGAAAGGGTGACATTTAT  
TTTGACAAGTTTTAGGCATCAGCTGGCATCAGTGTGTTTCAACTCCATTATTTGAAGTGAAATCCTCAC  
CTGGGGTCTCTGTGTGCAAAGCTGTCTTTTGAAGAACAGTTTGGTTGATGTCATGCCTTAGTAGCCAAA  
ATGCTACACTCTAGACTTACAAGTGGGAGTTAAGAGAGGTCTGGAAGTGTCCAACAAGGAATTCACACC  
TCTGCCTCTTTGACCAAGAAACAACTTTACACAGTTGGTAAGTGGGTCCATAACTGGCAGGATTTTAAAT  
TGTATTTTGCTCAAATCTATGGGAACAAAAGTCAAGGTATCACTACCTAGAAGTAATGATATACAGTTTT  
CTTCTAGTGGCTTGAATCTGGACTTCCTCAATTATTATTCACATTTTCTCTCTTATAGGTTTTCTGT  
TTTCTACTTTCTTTTTCTTATCTGTGTTTCCCTTTCTTTGTTGGCTCATTAACTTTTGACTGAAT  
TACAATTACTCCTTTTATTAAAGTCCATATTATGTGAATCATTTCCATGAAAATTTCTAAGAAAACCTCC  
AAACTCTCTAAATAGTAGCTAACTTTTATTTTTTAAATGAGTCGTGGGGTAGTGCTTCACTTGAGAT  
GCTTTGAAAGAGCCCTAAACATTGGGAACCACTTCACTAATTTGGAGACATTTCTACTGGTTGTGACTA  
CCCCCTTATGATCCTTCACTTCAATTTATGTCCCTAAACATCACAATGTAATATCATTTTTGTAGTTT  
CAGCTCACAGAGATTCTTACACTTGGGGTAAACACTATCCATGCATTACTTACTGGTAATTACCTGCT  
GGTATATAATTCCATGTAGCCTTTAATATGCTGGGTATCAAATCTGTTCACTGAGTTATGACCAGATA  
AATAATAGATATGCACATGAAAGATGCAAACTTGTGTGATTATTAAGCCAGCCATGCAGGTCCATGATA  
GAAACAGCAGGTGATGACTCTGCACTCTCATTGTCAAGGTTAGCTATATCCCAGTTGCAAAACAGCCAG  
ACTTGAGCTGTGCTCTGGTCACTTTGAGTTTAAAGCCCTTTGTTGTATAAGGCTGTGGAAAGTTGTA  
CAATGGCTGAAGCCATGTTGTTAATATGGCTGATGGGAGCATCCCTGCAGCTGAACCCAGCACTTTTAT  
GCTCCCACTGTGGTTGAGCTTTATGTTTACAGTCTCAGCAACAACACTTATGCATCCAAACACTCACAAA  
TGAAACCTGAAAGAATCTTTTCTGAGCCTCTTAAAGAGGAAATGATGATAACATTAAAGACTCTGAAC  
ACCCAAGGTTGGTGTACATATAAAAAATTAAGCTGATGACTTTGCAGTGACTCAAGTTGTCTCTTTATCA  
TGGTTTACCAGGTAGAGTGCTGGCTATTACTATATAATGAAGCCCACTGGCTTGACTTGTAAGTTCAAC  
CTAAACCACAATCCTAGACCATCATGGATTTAGGAGTAGATTCTTCTGAAATCCCACATCCAGAACTA  
GACATTAGAATGTTGAGGCAGTTTCCAGAGAAACAAGCATATTGCCTCATGGATGAAAGACTTGTAAGTT  
CTAGTTTCACTGACTTGTATATCTACTTACATACAACAGGGAGGCAAGAGGATTCTCTGTCTCTCTGG  
TGACTGAGTGTAAGTATGTGCAAGTCTGCAGCACAGTGACCAAACTGACAATCGAGCTCTGGATCAC  
CACTTGATTATGATAGACTCATTTTAAAGCAGCTTAGGAACTAATTAACATGGAGGATGAATTACC  
TTCTATCCCTTGAGATAAGACATCTTTCAAGTTTCAATTAAGGATTGTTGCTGTTTTATAGTTACTCT  
GTTTCATCAGTGTAATGGTGATGCGTGTCTAGGTGTGCAGCTATTTGAGGGACTAAGGGATGGAGAT  
ATTCTGTCAAATGAATCTCTTCAATATACCAGTTTGTGGGAGGGATATGAGACATGTGGATGGCAGTGAG

FIGURE 26.2

63/64

AGATCGTGCCTCTAGATCTTGATGGAGGCTTGGTGAGACACACTTAAATAAGCACGTGGAGGTTAGAATA  
GAGGGCAGAGTAAAAGGAAGCTCCATCTGAGCAAGTACACCAAATGATCTCAGCCCTGCAACTTGACCCA  
GGTAGGGCCACCACTACGCCTTCACTTGTCAACCAAGCTCCAACCACAGAGAGTTTGACAAGTTTGTGTT  
ATGATGTTGGCTTGGCTTTGTATTTTAAATTAACCTTGGATTTTGTAGTGGTTTGTGCATATAACTGTCTG  
AGTTTGGTAGGTAGGATTACTTTGAAAAGGGTTTACTAGTGTGGTCCTCCGGGTAGAATTTAGCTGTAAC  
ATGTTGTTAGCCAGCCTGTAGACTGTTAATTACTTAATAATCTCATTGGGAAAAATACTAGTAGTTTTATA  
TTTGGATGACATAATTGAAAAAGCAGATTAGCTGCTACTACTTTTAAAAGACTTAAGGTCGGGATGCCT  
TTTTTCCATGTAAGGAAATGAAAAGACCCAAAATCTTCAGGCAAAAAGCAAGTTGCAAAATTAGAAACC  
ATTGGCTAAAAATGTGTTTTGTTGAGTTTCCAAATGGATGAATTTTCATTGGACATTACATCACTAAAT  
TCATTAGATTTTGTCTGCATTGGAAAGATACTCTTCTAGCATATCTTTCCCAAAGATATCTAATTTGGAT  
TCTGTTTCATGCAAAATTGCATCCCGGAGGTTGAAGTTGGAGTTTGGAGTTGGAAAATATCTTTGAAGGC  
AGAACTCAGTTGATGTTGAGGGTGAAGCCTCACATACTTCTCAACAGACATGATAAAATTCACCTGCATG  
AGTTGGCAGGTGGGAGAACCAAACCTGGATCACTGGGTAAGACTACTCAGTAAAGCAATGAACCTGCTTGT  
TAGAGAAGCATCACTATCCCCATTGAGAAAAATGTGTGGCAAGATGATACAGCTACACAGTATCAAATGA  
ATGGGTCAATTCAGCACCCCAAAATTTAATCTGTGGGGAAAAATTATTGAGCCAGTTGTGAGTGTCTG  
TTACATGACTGGCAGACTAAATCTTCATCGTTGTGTTATGTTGTGTTGTTCTCATTTTCACTCGC  
ACGGCCTTATTCTCATAATTAATCTAATTCATTTTCTTTAGTGTTAGTAGACTCCAACAACAGAAG  
TGGCATCTGTGTATTCTAATCAGCATTTACCCTGGCAGGAGACTAATCAGATAGGCCGGTCTCAGACAT  
TAATCCTACCATTGTGATTTTGGTGAAGGAAAAAGTATTAATCTCTTTCCATCCTCCTCAGAAA  
TATAGAAGCCCTCTTTACCAAATCATCACATTTTACTCTGTAATCTACCAGCTAAAAGAAAAATTGCATT  
GAAGCCCCACAAAGCCAGATTGCAGTTCTTGCCCTTTTTCGCTCTGACATGAGATGTTAAAGAATTATT  
CATTGTGCTCACATTGGGTTAGGGGACACTGAACTGCTTTTAGATCCATGATCAGTCATCATTCTCTA  
AGAGATTGGAGCTTTGCTGTTTCAATTAACGTGTCAGTGTAGACTAATGGTGTTAATAAAAAATCAATCAA  
AATTTCAAACCTCTTTTGCCAGTGACCTCAATTTTGTGGCTCTGTGATTTGTATCAGACTTTGAGGAGGG  
AAGGGGAAGTGAAGGAAGCCTACGTCCAGGCCCTGACAGGATGCTGCAGTAGCAAGCTCAAGCTCGCC  
TGCTTGCCAGCATTTGCTGGTGAGCAGCAGCATGCAGACCAGCTGTGGGAAGCCTCCTGAAGAATGCCCC  
AGCTGATGCTTTAGCTGGGAATAGTTTGTCTTATTGGGGAACCTCATTGTTCTCCAGTCTCTGCAGCAG  
GAAGCCAGCTGTCTATTCGGAGGGAATTCAGATGCTTTACCTTTTGGTTTTGTCTGCATCACTCAT  
GTGGCTACGAAAGTGTCTCTGAGAATAGAGCCCAATGTGGTGACAATGGGTAGTCAAAATGCACCCAGAT  
GCTCAAGCCCTGTTGTGGTTCTGCAGTGTATGAAATGGGAGGAAGGAGACCCTGGACAGTAAGCAAA  
ATTGGAGACACTCCAACGAGGCTAAGTTAATGCCGTGTTGCCAGAAACAAGATCTAGCTTCTCATTTGGT  
CAGCCTAGCATGCAACCAGTGGTGTCTGGTAAAATGTTTAAACAACCAGCTCGCTGAGAATAGAAAGCAC  
CTGTTTGCACCAATTGCCAAATTTCCATGGCATAAATACTACCACTTTAGATGATTTTAAAGCTACCAAT  
GTGATGTCAGTGAACACATGGTTGGAAAGAGATGCACGCAGTTGGCTCTTGCAAGCCTGGGCAAAAATGC  
TTCAACACGCCACTGGATGCAGCCAGTCAGAGGGTTCATATTTAATATATGTGTTCTATGTGGACACACAC  
AGACACACACACAAACTCACCTTTACACACACACTTCGATGACTAAAACAATTACATAGTTTTAAGAT  
ATGAATCAATGTGTGAATGTAGAAAGCTTATGATAAGGCCCTAGAGGTATGGGTTGCCCTGGAAGCCTAG  
GTTTTAAGCAGGAGAATAGCTGAGAAGAATGAAGCCCTCCTGAGCTGAAAGGAGAGATGGATCAATGGAG  
ATGGTTCCATCATCTCCTTCCATATCTCACAGGTAAAATGGGCACTCAGAAAACCTCACGATTGATTTT  
TAAAAAGATAAGTGAGTGTTTTTATTTTATTATTATGTCATCATTATTTGATTTACAAATGCTATT  
TGTAACCTTTTACATGTAACTAGGATAAAGTATTTACGGGAACCTATGGAGAATAGCACAATCCAGAATT  
TACTGTGTTTTTCTTTATGTGACGTGAAACTCAGTAATTCTCCACCTTCACATTGTTGTTTATAAGA  
ATTTTACTTTAGTTATTAGGGAATCTAAGTTTTTGTAAACATTTGTTTTTAGTTAAAAGTATCTACTTA  
CTGTTTTAGCTCTGAACTCAAACCAGAAATATCTCTGTATCAATTGCATGACTATTACAGAAACAATAATCC  
AAACCAAAATAATTTCTTTTCCACCCAGTACGAAGAAAATAAGCTCAGTAACAAGAGGCATAAACTAA  
AGTATATAATGAGGCTTTCAATAATACACACACACACACTCACACACACACATACACTTTTTAA  
TTTTAAATTAGGCCTCCACACATAAATCATTTTGAAAGTAGAATAGAAAATCTCAAAGAATTCATTCTC  
CTGGTCTGTGTCATCTTCTGCAGTTAATAAGAGGTTTGTATCTGGAAAGATGGAAGAACTTGTTCTAAAA  
TCTTATTTTTCAAAAAAAATTTCCATTTTCTCTCTGGGCTGTATCCATGGTTGAATGTTAGCCCTGGA  
GGAGATCCATGTCTTACTCGCTCTTTCTGGCCCTTCTGTCTTTTGCCTCTGCAATTTCTTTTGTAGCTGG  
CACGATAGCAGGAGCTGGGGTCTATCTTTTCATGGTATGCTACAATATTGTCTTACTGGAAAATGG  
TAACATCCGGGTCTGATTTAATTGGCATTACACTTACACAGGAGCTCTGAGCACCCCGCTCACCACACCA

FIGURE 26.3



**FIGURE 26.4**

I

LISTE DE SEQUENCES

<110> INSERM  
ALLIEL, Patrick  
PERIN, Jean-Pierre  
RIEGER, Francois

<120> FAMILLE DE SEQUENCES NUCLEIQUES ET DE SEQUENCES  
PROTEIQUES DEDUITES PRESENTANT DES MOTIFS RETROVIRAUX  
ENDOGENES HUMAINS, LEURS SEQUENCES FLANQUANTES ET LEURS  
APPLICATIONS.

<130> 598EXT21

<140>  
<141>

<150> 9807920  
<151> 1998-06-23

<160> 122

<170> PatentIn Ver. 2.1

<210> 1  
<211> 2599  
<212> ADN  
<213> Homo sapiens

<400> 1  
atccccctgcc ttaatcgcca agctccttca ggagaacaaa gaacaggcca ttaccctgga 60  
gaagactggc aactgatttt acccacaagc ccaaacctca gggatttcag tatctactag 120  
tctgggtaga tactttcacg ggttgggcag aggccttccc ctgtaggaca gaaaaggccc 180  
aagaggtaat aaaggcacta gttcatgaaa taattcccag attcggactt ccccgaggct 240  
tacagagtga caatagccct gctttccagg ccacagtaac ccaggagta tcccaggcgt 300  
taggtatacg atatcactta cactgcgcct gaaggccaca gtccctcaggg aaggctcgaga 360  
aaatgaatga aacactcaaa ggacatctaa aaaagcaaac ccaggaaacc cacctcacat 420  
ggcctgctct gttgcctata gccttaaaaa gaatctgcaa cttccccaa aaagcaggac 480  
ttagcccata cgaaatgctg tatggaaggc ccttcataac caatgacctt gtgcttgacc 540  
caagacagcc aacttagttg cagacatcac ctcccttagcc aaatatcaac aagttcttaa 600  
aacattacaa ggaacctatc cctgagaaga gggaaaagaa ctattccacc cttgtgacat 660  
ggtattagtc aagtccttc cctctaattc ccatcccta gatacatcct gggaaggacc 720  
ctacccagtc attttatcta ccccaactgc ggttaaagt gctggagtgg agtcttggat 780  
acatcacact tgagtcaaat cctggatact gccaaaaggaa cctgaaaatc caggagacaa 840  
cgctagctat tcctgtgaac ctctagagga tttgcgcctg ctcttcaaac aacaaccagg 900  
aggaaagtaa ctaaaatcat aaatcccat ggccctccct tatcatattt ttctctttac 960  
tgttctttta ccctctttca ctctcactgc accccctcca tgcgcgtgta tgaccagtag 1020  
ctccccttac caagagtttc tatggagaat gcagcgtccc ggaatatattg atgccccatc 1080  
gtataggagt ctttctaagg gaacccccac ctctactgcc cacaccata tgccccgcaa 1140  
ctgctatcac tctgccactc tttgcatgca tgcaaatact cattattgga caggaaaaat 1200  
gattaatcct agttgtcctg gaggacttgg agtcaactgc tgttggaactt acttcaccca 1260  
aactggtagt tctgatggg gtggagtcca agatcaggca agagaaaaac atgtaaaaga 1320  
agttaatctc caactcacc gggtagatgg cactctatgc ccctacaaag gactagatct 1380  
ctcaaaacta catgaacccc tccgtaccga tactcgctg gtaagcctat ttaataccac 1440  
cctcactggg ctccatgagg tctcgcccca aaaccctact aactgttggga tatgcctccc 1500  
cctgaacttc aggccatag tttcaatccc tgtacctgaa caatggaaca acttcagcac 1560  
agaaataaac accacttccg ttttagtagg acctcttgtt tccaatctgg aaataaccca 1620  
tacctcaaac ctacactgtg taaaatttag caatactaca tacacaacca actcccaatg 1680

2

```

catcaggtgg gtaactcctc ccacacaaat agtctgccta cccctcaggaa tattttttgt 1740
ctgtgggtacc tcagcctatc gttgtttgaa tggctcttca gaatctatgt gcttcctctc 1800
attcttagtg cccctatga ccatctacac tgaacaagat ttatacagtt atgtcataac 1860
taagccccgc aacaaaagag taccattctt tccttttgtt ataggagcag gagtgcctag 1920
tgcactaggt actggcattg gcggtatcac aacctctact cagttctact acaaactatc 1980
tcaagaacta aatggggaca tggaaacggg cgccgactcc ctggtcacc tgcaagatca 2040
acttaactcc ctagcagcag tagtccttca aaatcgaaga gcttttagact tgctaaccgc 2100
tgaaagaggg ggaacctgtt tatttttagg ggaagaatgc tgttattatg ttaatcaatc 2160
cggaatcgtc actgagaaag ttaaagaaat tcgagatcga atacaacgta gagcagagga 2220
gcttcgaaac actggaccct ggggcctcct cagccaatgg atgccctgga ttctccctt 2280
cttaggacct ctagcagcta taatattgct actcctcttt ggaccctgta tctttaacct 2340
ccttgtaaac tttgtctctt ccagaatcga agctgtaaaa ctacaaatgg agcccaagat 2400
gcagtcgaag actaagatct accgcagacc cctggaccgg cctgctagcc cagcatctga 2460
tggttaatgac atcaaaaggca cccctcctga ggaatctca gctgcacaa ctctactacg 2520
ccccaattca gcaggaagca gtttagagcg tctcgcccaa cctccccaac agcacttagg 2580
tttctctgtt gagatgggg 2599

```

```

<210> 2
<211> 1326
<212> ADN
<213> Homo sapiens

```

```

<400> 2
gccgcctggc actcctgagg gaagtataaa ttataacacc atcttacagc tagacctctt 60
ttgtagaaaa ggcaaatgga gtgaagtgcc ataagtacaa actttctttt cattaagaga 120
caactacaaa ttatgtaaaa agtgtgatgt atgccctaca ggaagccttc agagtctacc 180
tccctatccc agcatccccg actccttccc caactaataa ggacccccct tcaacccaaa 240
tggtccaaaa ggagatagac aaaagggtaa acagtgaacc aaagagtgc aatattcccc 300
aattatgacc cctccaagca gtgggaggaa gagaattcgg ccagccaga gtgcatgtgc 360
ctttttctct ccagactta aagcaataa aaacagactt aggtaaattc tcagataacc 420
ctgatggcta tattgatgtt ttacaagggt taggacaatt cttgatctg acatggagag 480
atataatgtc actgctaaat cagacactaa ccccaaatag gagaagtgc accataactg 540
cagcctgaga gtttggcgat ctctgtatc tcagttagt caatgatagg atgacaacag 600
aggaagaga atgattcccc acaggccagc aggcagttcc cagtctagac cctcattggg 660
acacagaatc agaacatgga gatttgtgct gcagacattt gctaactgt gtgctagaag 720
gactaaggaa aactaggaag aagtctatga attactcaat gatgtccacc ataacacagg 780
gaagggaaga aaatcctact gcctttctgg agagactaag ggaggcattg aggaagcgtg 840
cctctctgtc acctgactct tctgaaggcc aactaatctt aaagcgtaag tttatcactc 900
agtcagctgc agacattaga aaaaaacttc aaaagtctgc cgtaggcccc gagcaaaact 960
tagaaaccct attgaacttg gcaacctcgg tttttataa tagagatcag gaggagcagg 1020
cggaacagga caaacgggat taaaaaaaag gccaccgctt tagtcatgac cctcaggcaa 1080
gtggactttg gaggtctctg aaaagggaaa agctgggcaa attgaatgcc taatagggct 1140
tgcttccagt gcggtctaca aggacacttt aaaaaagatt gtccaagtag aagtaagccg 1200
ccccctctgc catgcccctt atttcaaggg aatcactgga aggccactg cccaggggga 1260
caaaggctct ctgagtcaga agccactaac cagatgatcc agcagcagga ctgaggggtc 1320
ctggggg 1326

```

```

<210> 3
<211> 10499
<212> ADN
<213> Homo sapiens

```

```

<400> 3
ccctggggcg ggcttccttt ctgggatgag ggcaaacgc ctggagatac agcaattatc 60
ttgcaactga gagacaggac tagctggatt tcctaggccg actaagaatc cctaagccta 120
gctgggaagg tgaccacgtc cacctttaaa cacggggctt gcaacttagc tcacacctga 180

```

```

ccaatcagag agctcactaa aatgctaatt aggcaaagac aggaggtaaa gaaatagcca 240
atcatctatt gcctgagagc acagcaggag ggacaacaat cgggatataa acccaggcat 300
tcgagctggc aacagcagcc cccctttggg tcccttccct ttgtatggga gctgttttca 360
tgctatttca ctctattaaa tcttgcaact gcactcttct ggtccatggt tcttacggct 420
cgagctgagc ttttgctcac cgtccaccac tgcgttttgc caccaccgca gacctgccgc 480
tgactcccat ccccttggtat cctgcagggt gtccgctgtg ctccctgatcc agcgaggcgc 540
ccattgccgc tcccaattgg gctaaaggct tgccattgtt cctgcacggc taagtgcctg 600
ggtttgttct aattgagctg aacactagtc actgggttcc atggttctct tctgtgacct 660
acggcttcta atagaactat aacacttacc acatggccca agattccatt ccttggaaac 720
cgtgaggcca agaactccag gtcagagaat acgaggcttg ccaccatctt ggaagcggcc 780
tgctaccatc ttggaagtgg ttaccacca tcttgggagc tctgtgagca aggaccccc 840
ggtaacattt tggcaaccac gaacggacat ccaaagtgtt gagtaatat ggaccacttt 900
cacttgctat tctgtcctat ccttccttag aattggagga aaataccggg cacttgtcgg 960
ccagttaaaa acgattagtg tggcaccgg acttaagact caggtgtgag gctatctggg 1020
gaagggtttt ctaacaaccc ccaacccttc tgggttgggg acttggtttg cctcaagcca 1080
gcttccactt tcagttttct tggggaagcc gagggccgac tagaggcaga aagctgtcgt 1140
cctgaactcc cggcagtagc cgggtgagat catggtgtag ccagaagtct caacagtcgc 1200
ccatgcatgc ctcccattt ttcccttctga cccataacct ctgggtccca accacaacct 1260
tcttcaaagt gtagcccaa aattctcctt acctctgaat atacttcttc tgatccctgc 1320
ctcctaggtt ctattggttc agacttccat ttcctctagc aagttgtatc tccaaaggga 1380
tctaaggagc ctctgcgctg cgtccttagg cacctaggct ataaccagg gagtcttctc 1440
cctgggtgctt ctcccattt aggcatacag ctcttgacat gggcagttat gtaggacca 1500
ctccccacca cccttgccag ggccccaaat ttgtaaatgg ctgagggaaa agagagacag 1560
aggagagaga gagaaatgga ggagaaagag agagagacag agaggagaga gagacagtga 1620
gagagacaga agagagagag agacaaagag gagagagaga gagtcaaaga gagaagaaa 1680
gagaaagaaa tagtaaaaaa cagtgtgccc tattccttta aaagccaggg taaattttta 1740
acctgtactt gataattgaa ggtcttctct gtgacctat agcactccaa tccactttgt 1800
ggtcagtgta aataagagca taggcccgaat gcactgaggc cattgacaac ccgtagcttc 1860
cctatcaaaa atccttaacc cagtaaccgg cagatggacc aaatgcattc agtcggtagc 1920
gcaactgctt tgctaaaagt agaaaaagta cttttagagg aaacctcatt gtgagcaca 1980
ctcacctgtt cagaattatt ctaataaaaa aagcaaaaag gtagcttact aactcaaaaa 2040
tcttaaaagt tggggctatt ctgttagaaa aaggtaatgt aactccaacc actgataatt 2100
cccttaaccc agcagatttc ctaacgggat ttaaatctta attaccatac aaaggctccg 2160
ccagacctag gcggaactcc cttcaggaca ggacgataga tggttcctcc caggtgattg 2220
aggaaaaaaa ccacaatggg tattcagtaa ttgatacggg gactcttgtg gaagcagagt 2280
tagaaaaatt gcctaataac tgggtctctc aaacgtgtga gctgtttgca ctacgccaag 2340
ccttaaaagt cttacagaat caaaagacta tctcaatcct gattcaaaag gttagctaca 2400
ccctctctgt aatgcatttg cataagaact tgtttatggg aatgcattct gatggggcag 2460
ctgggttgtt ataaaatagg aacccagccc agctctagga ctacccctg agcgcaaaag 2520
caatgttggg catgctggtt aaggaccact agaatccagc agcccagacc cctttctttg 2580
tggctcaagaa aggcgggaaa aggggtgcag gactgctaca tcggtaagca taactaatcc 2640
gataaacaga ggtccatggg tggttacgca ccctggaaag gaactcaccc ctgagcaca 2700
aggcaatgtt gggcacgctg gtaaaaggacc actagaatcc agcagcctgg acccctttct 2760
ttgtggtcaa gagaggcagg aaaacagggt caggactgca acatcagtga gcataactaa 2820
ttcgataagc agagggtccat gggtggtgat gcaccctgga aagaataagc attaggacca 2880
tagaggacac tccaggacta aagctcatcg gaaaatgact agggttgctg gcatccctat 2940
gttctttttt cagatgggaa acgttccccg caagacaaaa acgcccctaa gacgtattct 3000
ggagaatttg gaccaatttg accctcagac actaagaaa agaacgactt tattcttctg 3060
cagtgccggc tggcactcct gagggaaagta taaattataa caccatctta cagctagacc 3120
tctttttagt aaaaggcaaa tggagtgaag tgccataagt acaaaacttc ttttcattaa 3180
gagacaactc acaattatgt aaaaagtgtg atttatgccc tacaggaaag cttcagagtc 3240
tacctcccta tcccagcatc cccgactcct tccccacta ataaggacc ccctcaacc 3300
caaatggtcc aaaaggagat agacaaaagg gtaaacagtg aaccaaagag tgccaattat 3360
ccccattat gacccctcca agcagtgagg ggaagagaat tcggcccagc cagagtgcac 3420
gtgctttttt ctctccaga cttaaagcaa ataaaacag acttaggtaa attctcagat 3480
aaccctgatg gctatattga tgttttacia gggtaggac aattctttga tctgacatgg 3540
agagatataa tgtcactgct aaatcagaca ctaaccctaa atgagagaag tgccaccata 3600
actgcagcct gagagtttgg cgatctctgg tatctcagtc aggtcaatga taggatgaca 3660

```

acagaggaaa	gagaatgatt	ccccacaggc	cagcaggcag	ttcccagtc	agacccctca	3720
tgggacacag	aatcagaaca	tggagattgg	tgctgcagac	atttgctaac	ttgtgtgcta	3750
gaaggactaa	ggaaaactag	gaagaagtct	atgaattact	caatgatgtc	caccataaca	3840
cagggaaggg	aagaaaatcc	tactgccttt	ctggagagac	taagggaggg	attgaggaag	3900
cgtgcctctc	tgtcacctga	ctctcttgaa	ggccaactaa	tcttaaaagg	taagtttat	3960
actcagtcag	ctgcagacat	tagaaaaaaa	cttcaaaagt	ctgccgtagg	cccggagcaa	4020
aacttagaaa	ccctattgaa	cttggcaacc	tcggtttttt	ataatagaga	tcaggaggag	4080
caggcggaac	aggacaaacg	ggattaaaaa	aaaggccacc	gcttttagtca	tgacccctcag	4140
gcaagtggac	tttggaggct	ctggaaaagg	gaaaagctgg	gcaaattgaa	tgccctaata	4200
ggcttgcttc	cagtgcggtc	tacaaggaca	ctttaaaaaa	gattgtccaa	gtagaagtaa	4260
gccgccccct	cgtccatgcc	ccttatttca	agggaaatcac	tggaaggccc	actgccccag	4320
gggacaaaag	tcctctgagt	cagaagccac	taaccagatg	atccagcagc	aggactgagg	4380
gtgcctgggg	caagcgccat	cccatgccat	caccctcaca	gagccctggg	tatgcttgac	4440
cattgagggc	caggagggtg	tctcctggac	actggtgctg	tcttcttagt	cttactcttc	4500
tgccccggac	aactgtcctc	cagatctgtc	actatctgag	ggggtcctaa	gacgggcag	4560
cactagatac	ttctcccagc	cactaagtta	tgactgggga	gctttatttc	tttcacatgc	4620
ttttctaatt	atgcttgaaa	gccccactac	cttggttaggg	agagacattc	tagcaaaaga	4680
agggggccatt	atacacctga	acataggaga	aggaacaccc	gtttgttgtc	ccctgcttga	4740
ggaaggaaat	aatcctgaag	tctgggcaac	agaaggacaa	tatggacgag	caaagaatgc	4800
ccgtctcgtt	caagttaaac	ttaaaggattc	cacctccttc	ccctaccaaa	ggcagtagcc	4860
cctcagaccc	aaggcccaac	aaggactcca	aaagattgtt	aaggacctaa	aagcccaagg	4920
cctagtaaaa	ccatgcagta	acccctgcag	tactccaatt	ttaggagtag	agaaacccaa	4980
cagacagtg	aggttagtgc	aagatctcag	gattatcaat	gaggctgttg	ttcctctata	5040
gccagctgta	cctagccctt	atactctgct	ttcccaataa	ccagagggaag	cagagtgggt	5100
tacagtctgt	gaccttcagg	atgccttctt	ctgcatccct	gtacatccct	actctcaatt	5160
cttgtttgcc	tttgaagata	cttcaaaccc	aacatctcaa	ctcacctgga	ctattttacc	5220
ccaagggttc	agggatagtc	cccattctatt	tggccaggca	ttagcccaag	acttgagcca	5280
atcctcatat	ctggacactt	gtccttcggt	aggtggatga	tttacttttg	gccgcccaat	5340
cagaaaacct	gtgccatcaa	gccaccaca	cgctcttcaa	tttctctgct	acctgtggct	5400
acatggtttc	caaaccaaa	gctcaactct	gctcacagca	ggttacttag	ggctaaaaat	5460
atccaaaggc	accaggggcc	tcagtggagg	acacatccag	cctatactgg	cttatcctca	5520
tcccaaaacc	ctaaagcaac	taaggggatt	ccttggcgta	ataggtttct	gccgaaaaat	5580
gattcccagg	tatggcgaaa	tagccaggtc	attaaataca	ctaattaagg	aaactcagaa	5640
agcccaatac	catttagtaa	gatggacaac	tgaagtagaa	gtggctttcc	aggccctaac	5700
ccaagcccca	gtgttaagtt	tgccaacagg	gcaagacttt	tcttcataat	tcacagaaaa	5760
aacaggaata	gctctaggag	tccttacaca	gatccgaggg	atgagcttgc	aacctgtggc	5820
atacctgact	aaggaaattg	atgtagtggc	aaagggttga	cctcattgtt	tacgggtagt	5880
ggtggcagta	gcagctctag	tatctgaagc	agttaaaata	atacagggaa	gagatcttac	5940
tgtgtggaca	tctcatgatg	tgaatggcat	actcactgct	aaaggagact	tgtggctgtc	6000
agacaaactgt	ttacttaaat	gtcaggctct	attacttgaa	gggccagtgc	tgcgactgtg	6060
cactttgtga	actcttaacc	cagccacatt	tcttccagac	aatgaagaaa	agataaaaac	6120
taactgtcaa	caagtaattt	ctcaaaccta	tgccactcga	ggggaccttt	tagaggttcc	6180
tttgactgat	cccgaacctca	acttgataac	tgatggaagt	tcctttgtag	aaaaaggact	6240
tcgaaaagtg	gggtatgcag	tggtcagtga	taatggaata	cttgaaagta	atccccctcac	6300
tccaggaact	agtgtctcag	tagcagaact	aatagccctc	acttgggcac	tagaattagg	6360
agaagaaaaa	agggcaataa	tatatacaga	ctctaaatat	gcttacctag	tcctccatgc	6420
ccatgcagca	atatggaaag	aaagggaatt	cctaacttct	gagagaacac	ctatcaaaac	6480
tcaggaagcc	attaggaat	tattattggc	tgtacagaaa	cctaaagagg	tggcagctct	6540
acactgccgg	ggtcatcaga	aaggaaagga	aagggaataa	gaagagaact	gccaaagcga	6600
tattgaagcc	aaaagagctg	caaggcagga	ccctccatta	gaaatgctta	taaaaacaac	6660
cctagtatat	ggtaatcccc	tccgggaaac	caagccccag	tactcagcag	gagaaacaga	6720
atggggaacc	tcacgaggac	agttttctcc	cctcgggacg	gctagccact	gaagaaggga	6780
aaatactttt	gcctgcaact	atccaatgga	aattacttaa	aaccttctat	caaacctttc	6840
acttaggcct	cgatagcacc	catcagatgg	ccaaatcatt	atttactgga	ccaggccttt	6900
tcaaaaactat	caagcagata	gtcagggcct	gtgaagtgtg	ccagagaaat	aatccccctgc	6960
cttatcgcca	agctccttca	ggagaacaaa	gaacaggcca	ttaccctgga	gaagactggc	7020
aactgatttt	accacacaag	ccaaacctca	gggatttcag	tatctactag	tctgggtaga	7080
tactttcacg	ggttgggcag	aggccttccc	ctgtaggaca	gaaaaggccc	aagaggtaat	7140

aaaggcacta	gttcatgaaa	taattcccag	attcggactt	ccccgaggt	tacagagtga	7200
caatagccct	gctttccagg	ccacagtaac	ccaggagta	tcccaggt	taggtatacg	7260
atatcactta	cactgcgcct	gaaggccaca	gtcctcaggg	aaggtcgaga	aaatgaatga	7320
aacactcaaa	ggacatctaa	aaaagcaaac	ccaggaaacc	cacctcaca	ggcctgctc	7380
gttgccctata	gccttaaaaa	gaatctgcaa	ctttccccaa	aaagcaggac	ttagcccata	7440
cgaaatgctg	tatggaaggc	ccttcataac	caatgacctt	gtccttgacc	caagacagcc	7500
aacttagttg	cagacatcac	ctccttagcc	aaatatcaac	aagttcttaa	aacattacaa	7560
ggaacctatc	cctgagaaga	gggaaaagaa	ctattccacc	cttgtagaca	ggtattagtc	7620
aagtcccttc	cctctaattc	cccattcccta	gatacatcct	gggaaggacc	ctacccagtc	7680
attttatcta	ccccaaactgc	ggttaaagtg	gctggagtgg	agtcttgga	acatcacact	7740
tgagtcaaat	cctggatact	gccaaaaggaa	cctgaaaaatc	caggagacaa	cgctagctat	7800
tctgtgaac	ctctagagga	tttgcgcctg	ctcttcaaac	aacaaccagg	aggaaagtaa	7860
ctaaaaatca	aaatccccat	ggccctccct	tatcatattt	ttctctttaa	tgttctttaa	7920
ccctcttcta	ctctcactgc	acccctccca	tgccgctgta	tgaccagtag	ctcccttacc	7980
caagagtttc	tatggagaat	gcagcgtccc	ggaaatattg	atgccccatc	gtataggagt	8040
ctttctaaagg	gaacccccac	cttcactgcc	cacaccata	tgccccgcaa	ctgctatcac	8100
tctgccactc	tttgcattga	tgcaataact	cattattgga	caggaaaaat	gattaatcct	8160
agttgtcctg	gaggacttgg	agtcactgtc	tgttggaactt	acttcaccca	aactgggtac	8220
tctgatgggg	gtggagtcca	agatcaggca	agagaaaaac	atgtaaaaga	agtaatctcc	8280
caactcacc	gggtacatgg	cacctctagc	ccctacaaag	gactagatct	ctcaaaaacta	8340
catgaaaccc	tccgtaccca	tactcgcctg	gtaagcctat	ttaataccac	cctcactggg	8400
ctccatgagg	tctcggccca	aaaccctact	aactgttgga	tatgcctccc	cctgaacttc	8460
aggccatag	tttcaatccc	tgtacctgaa	caatggaaca	acttcagcac	agaaataaac	8520
accacttccg	ttttagtagg	acctcttggt	tccaatctgg	aaataaccac	tacctcaaac	8580
ctcacctgtg	taaaatttag	caatactaca	tacacaacca	actcccaatg	catcaggtgg	8640
gttaactcctc	accacaaaat	agtcctgcct	ccctcaggaa	tattttttgt	ctgtgggtac	8700
tcagcctatc	gttgtttgaa	tggctcttca	gaatctatgt	gcttctctct	attcttagtg	8760
ccccctatga	ccatctacac	tgaacaagat	ttatacagtt	atgtcatatc	taagccccgc	8820
aacaaaagag	taccattctt	tccttttggt	ataggagcag	gagtgctagg	tgacttaggt	8880
actggcattg	gcggtatcac	aacctctact	cagttctact	acaaactatc	tcaagaacta	8940
aatggggaca	tggaaacggg	cgccgactcc	ctgggtcacct	tgcaagatca	acttaactcc	9000
ctagcagcag	tagtccttca	aaatcgaaga	gcttttagact	tgctaaccgc	tgaagagggg	9060
ggaacctgtt	tatttttagg	ggaagaatgc	tgttattatg	ttaatcaatc	cggaatcgtc	9120
actgagaaag	ttaaagaaat	tcgagatcga	atacaacgta	gagcagagga	gcttcgaaac	9180
actggaccct	ggggcctcct	cagccaatgg	atgccctgga	ttctccctct	cttaggacct	9240
ctagcagcta	taatattgct	actcctcttt	ggaccctgta	tccttaacct	ccttggtaac	9300
tttgtctctt	ccagaatcga	agctgtaaaa	ctacaaatgg	agcccaagat	cgagtccaag	9360
actaagatct	accgcagacc	cctggaccgg	cctgctagcc	cacgatctga	tgttaatgac	9420
atcaaaaggca	ccctcctga	ggaaatctca	gctgcacaac	ctctactacg	cccccaattca	9480
gcaggaagca	gttagagcgg	tctcggccaa	cctccccaac	agcacttagg	tttctctgtt	9540
gagatggggg	actgagagac	aggactagct	ggatttcccta	ggctgactaa	gaatccctaa	9600
gcctagctgg	gaaggtgacc	acatccacct	ttaaaccagg	ggcttgcaac	ttagctcaca	9660
cctgaccaat	cagagagctc	actaaaatgc	taattaggca	aagacaggag	gtaaagaaat	9720
agccaatcat	ctattgcctg	agagcacagc	aggagggaca	atgatcggga	tataaaccac	9780
agtcttctgag	ccggcaacgg	caacccccct	tgggtcccct	ccctttgtat	gggagctctg	9840
ttttcatgct	atttcactct	attaaatctt	gcaactgcac	tcttctgggt	catgtttctt	9900
acggcttgag	ctgagctttc	gctcgccatc	caccactgct	gtttgccgcc	accgcagacc	9960
cgccgctgac	tcccatccct	ctggatcatg	cagggtgtcc	gctgtgctcc	tgatccagcg	10020
aggcaccat	tgcgctccc	aatcgggcta	aaggcttgcc	attgttctct	catggctaag	10080
tgccctgggt	catcctaatt	gagctgaaca	ctagtcactg	ggttccatgg	ttctcttctg	10140
tgacccacag	cttctaatag	agctataaca	ctcacgcgat	ggcccaaggt	tccattctct	10200
gaatccataa	ggccaagaac	cccaggctcag	agaacacgag	gcttgccacc	atcttgggag	10260
ctctgtgagc	aaggaccccc	aagtaacaca	accatgaggg	tgcaaatgca	tgggccacta	10320
atggtagagc	aagaaaacag	aagggccctg	gttcctcgaa	ggcatcagtg	agctgaaatg	10380
cctgccttgg	attgtcctatt	cctaggtgtt	tttctgcctg	aagcagatta	aaccctttgt	10440
tcacttctcc	aagttagggct	tctattacag	cccaaatcaa	tccccacccc	agatgacat	10499

<210> 4  
 <211> 2784  
 <212> ADN  
 <213> Homo sapiens

<400> 4  
 ctccttcagg agaacaaaga acaggccact acccaagaga agactggcaa ctagatttta 60  
 cccatagtcc caaatctcag ggatttcagt atctactagt ttgggtagat actttcactg 120  
 gttgggcaga ggccttcccc ttaggacag aaaaggccca agaggtaata aacgttcacg 180  
 aaataattcc cagattcaga cttccccaag gcttacagag tgacaatggc cctgctttca 240  
 aggctacagt aacccaagga gtatcccagg tgttaggtat acaatatcac tcacactgag 300  
 cctggaggcc acagtcctca ggaaggtgg agaaaatgaa caaaacactc aaatgacatc 360  
 taaaaaagct aatccaggaa acccacctcg catggcctgc tctgttgccct atagccttac 420  
 taagaatccg aaactctccc caaaaagcag gacttagtcc atacaaaatg ctgtatggac 480  
 ggcccttccc aaccaatgaa cttgggcttg accgagagac agccaactta gttgcagaca 540  
 tcatctcctt agccaaatat caacaggttc ttaaaacatt acaggggagcc tgtccccaag 600  
 aagagggaaa ggaactattc caccctggtg acatgggtatt agtcaagtcc cttccctcta 660  
 attcccccac cctagatata tcctgggaag gaaactaccc agccatttta tctaccctaa 720  
 cggcagttaa agtggctgga gcggagtctt ggatacatca cactcaagtc aaacctgaga 780  
 tactgcaaaa ggaactcaaa aatccatgag acaatgctag ctattcctgt gaacctctag 840  
 aggatctgag cctgctcttc aaatgacaac cagggggaaa gtaactaaaa tcgtaaatcc 900  
 cctggccctc ccttatcata tttttctctt tactgttctc ttacccctt tcactctcac 960  
 tgcacccctg ccatgccact gcaccccgct catgccccgt ccatgccagt agctccctt 1020  
 agcaagagtt tctatggaga atgcagcgtc ccggaaatat tgatgcccc aatgtatagg 1080  
 gtttatctaa gggaaacccc accttactg cccacaccca tatgccccac aactgctata 1140  
 actctgccc acctttgcat catgcaata ctcattattg gacaggaaaa acgattaatc 1200  
 ccagttgtcc tggaggactt ggaggactca cttcactcat accagtatgt ctgattgggg 1260  
 tggagttcaa gatcaggcaa cagaaaaaca cataaaggaa gtaatctccc aactgacctg 1320  
 ggtacatagc acccctggcc cctacaaagg actagatctc tcaaaactac atgaaacct 1380  
 cctaccctac actggcctgg taagcctatt taataccacc ctgactgggc tccatgaggt 1440  
 ctgcggccaa aacctacta actgttggtg gtgcctcccc ctgcacttta ggccatacat 1500  
 ttcaatccct atacctgaac aatggaacaa cttcagcaca gaaataaaca ccacttctgt 1560  
 tttagtaggt cctctttcca atctggaaat aacccatacc tcaaacctca cctgtgtaaa 1620  
 atttagcaat actatagaca cagccaactc ccaatgcac aggtgggtaa ctcctccac 1680  
 acgaatagtc tgcctaccct caggaatatt tttgtctgt ggtacctcag cctatcattg 1740  
 tttgaatggc tcttcagaat ctgtgtgctt cctctcattc ttagtgcccc ctatgcccc 1800  
 ctacactgaa caagatttat acaatcatgt catacctaa ccccgcaaca aaagagtacc 1860  
 cattcttccc tttgttattg gagcaggagt gctaggcgga gtagctactg gcattggcgg 1920  
 tatcacaacc tctactcagt tctactacaa actgtctcaa gaactaaatg gtgacatgga 1980  
 atgggtcgtg gataccctgg tcaccttgca agatcaactt aactccctag cagcagtagt 2040  
 ccttcaaaat cgaagagctt tagacttgct aaccggcgaa agcgggggaa cctttttatt 2100  
 ttttagaggaa aaatgctgtt gttatgttaa tcaatccgga atcatcaccg agaaagttaa 2160  
 agaaattcaa ggtcgaatat aacgtagagc aaaggagctg caaaacactg gaccctgggg 2220  
 cctcctcagc caatggatgc cctggattct ccccttctta ggacctctag cagctataat 2280  
 attgttactc ctctttggac cctgtatctt taacctcctt gtttaagttt tcttttccag 2340  
 aatcgaagca gtaaaactac aaatcgttct tcaaatggag cccagatgc agtccatgag 2400  
 taaaatctac cacggacccc tggaccggcc tgctagccca tgctctgatg ttaatgacat 2460  
 caaaggcacc cctcccgagg aaatctcaac tgcacaacct ctactacgcc ccaattcagc 2520  
 aggaagcagt tagagtgggt gttggccaac ctccccaca gcagttgggt tttcctgtt 2580  
 agagggggga ctgagagaca ggaataacta gatttcttag accaactaag aatccctaag 2640  
 actagctggg aaggtgaccg ctccacctt taaacaccgg gcttgcaact tagctcacgc 2700  
 ccaaccaatc agatactaaa gagagctcac taaaatgcta attaggcaaa aacaggagat 2760  
 aaagaaatag ccaatcatct gttg 2784

<210> 5  
 <211> 1799  
 <212> ADN  
 <213> Homo sapiens

<400> 5  
 gggattctta gtcggcctag gaaatccagc taatcctgtc tctcagtcce cccactcaac 60  
 agggaaaccc aagtgcgtgtt ggggaggttg gctgacgacc agtctaactg cttcctgcgg 120  
 aattggggca tagtaggggt tgtgcagttg agatttcctc gggaggggtg cgttcgata 180  
 cattacaatt ggagcatggg ctatagggcc ggtccagggg tccacggtag atcttagtca 240  
 tggactcat ctgggggtcc atttgaagaa cgattttag ctttacaact ttgattctgg 300  
 aagagacaaa cttaacaagg aggttaaaga tacaggggcc aaagaggagt atcaatatta 360  
 gagctgctag agatcctaag aaggggagaa tccagggcat ccattggctg aggaggcccc 420  
 agggctctgtt gtttttgaag ctctctctgt ctacgttcta ttcaatctcg aatttcttca 480  
 actttctctg tgacaattca ggattgatta acataataac aacattcttc cgctaaaaata 540  
 acataataac aacattcttc cctaaaaat aaacagcttc cccctcttcc agagggttagc 600  
 aagtctaaag ctcttcaatt ttgaaggact actgatgcta ggaagttaa gttgatcttg 660  
 aaggtagacca gggagtcggc aacccattcc atgtcaccat tgagttcttg agatagtttg 720  
 tagtagaact gtagagagt tgtgtaccg ccaatgccag aacctagtcc acctagcac 780  
 cctgctccga taacaaaagg aagaatgagt actcttttgt tgtggggctt aggtacaaca 840  
 taattgtata aatcttgttc agtgtaaat gtcaggggg cactaagaa gagaggagc 900  
 acatagattc tgaagagcca ttcaaacac gataggctaa ggtaccacag acaaaaaata 960  
 ttcctgaggt taggcagact attcgtgttg gaggagttac ccacctgatg cattgggagt 1020  
 tgggtgtgtc tacagtattg ctaaaattta cacaggtgag gtttgaggta tgggttattt 1080  
 ccagatttga aacaagaggt cctactaaaa cggaagtggg gtttatttct gtgctgtagt 1140  
 tgttccattg ttcaggtaca gggattgaaa tgcattggcct gaaatacagg gggaggcaca 1200  
 accaacagtt gctaggggtt ttgaccgaga cctcatggag cccagttagg gtggtattaa 1260  
 ataggcttac caggcaagta tgggtatgga ggggttcatg tagttttaag agatctagtc 1320  
 cttttagggg gctaggggtg ctatgtaccc gggtcagttg ggaggttact tcctttacat 1380  
 gtttttctct tgcctgatct tgaactccac cccctcaga cataccagta tgggtgaagt 1440  
 aagtcgaca gacagtggct ccaagtcttc caggacaact aggattaatc attttccctg 1500  
 tccaataatg agtatattgca tgcattgcaa gaggggcaga gttatagcag ttgtggggca 1560  
 tatgggtgtg ggcagtgaag gtggagtttc ctttaggtaa actcctattt gatggggcat 1620  
 caatatttct ggggaagccg attcttcata gaaactcttg gtaaggggag ctgctggttg 1680  
 tacagcagca tggagggggg gcagtgagag tgaaaggggg taagagaaca gtaaaagaaa 1740  
 aaatatgata agggagggcc atggggattt acgattttag ttactttcct cacggttgt 1799

<210> 6  
 <211> 1489  
 <212> ADN  
 <213> Homo sapiens

<400> 6  
 tgggtgcttg cccgggcact ctacgtcctg ctgctggatc atctggttag tggcttctga 60  
 ctacagaggac ctacgtcccc tggggcagtg ggccttacag tgattccctt gacacgaggt 120  
 gcatggacga gggggcggct tatttctatt tggacaatct tttttaaggt gtcctttagt 180  
 accgcactgg aagcaaaccc tattaggcat ttgatttgcc tagcttttcc cttttccagt 240  
 gcctccaaag tccgcttgcc tgaggggcat gactaaagcg gtggcctttt ttttatccca 300  
 tttgtcccat tctgcctgct catcctgatc tctattataa aaaactgagg ttgccaaagt 360  
 caatagggtt tctaagtttt gttccgggcc taaggcagac ttttgaagt ttttccctaa 420  
 gctctgagct gactgagtga taaactatc ctttaagatt agttggcctt cagtagagtc 480  
 agttgacaga gagaggtatg ctctctcaat gcctccgtta gtcactccag aaaggcggta 540  
 ggattttctt cctttccctg tgttatagtg gacatcattg aataactcac aggtcttctt 600  
 ctagttttcc ttagtcttcc tagcacgcaa gtttagcaat gctcgcggca ccaatctcca 660  
 tgttctgatt ctgtgtccca gtgagggctt acactgggaa ctgcctgctg gcctgtgggg 720  
 aatcgttctc tttcctctgt tgtcgacct ccatgacct gactgagata ccagagatcg 780  
 ccaaaacttc aggtgcagct tacggcgaca cttctgtcat ttggggtag tgtctgatt 840



```

agcagtaaca ttatatctct ccatatcaga tcaaaggatt gtcctaaacc ttgtaaaaca 900
tcaatatagc cattaggggt atctgagaat ttacctaggt ctattttaat ttaaagtctg 960
ggagagaaaa aggcacatgc actctggctg ggccgaattc tcttccccc actgcgtctg 1020
agagagaaaa aggtacgtgc actctggctg ggccgaattc tcttccccc gcttggaggg 1080
ggcataatcg gggaatattg gcattctttg gttagttgtt tacccccctg tctatctcct 1140
tttgaccgt ttgggttgaa ggggggtcct tattatttgg ggaaggagt c tgggggagtc 1200
tggggtaggg aggtagactc tgagggtctc ctgtagggca taaatcacac tttttacata 1260
attgcgagtt gtctcttaat gaaaagaaag tttgtacgta tgacacttca caccatttgc 1320
cttcttttct acaaaagagg tctagctgta agatggtgt ataatttatg cttccctcag 1380
gatgccaggt ttctccccct taaagagtat atcgttgcca ggcggtagct cagaagaata 1440
tgtctttttt ttcttagcat ctgagagtca aattggtccc aattctcca 1489

```

<210> 7  
 <211> 1216  
 <212> ADN  
 <213> Homo sapiens

```

<400> 7
taaagataca gggattgaaa tgtatggcct gaagtgcagg gtcatatagg tgggggtgg 60
gaaaatgggg tttccttttag aaaaactcct atacgatggg tcatcaata tccaggaag 120
ccgcattctc catagaagct cttggtaatg ggagctactg gtagtacagt ggcatggagg 180
gggtgcagtg agagtgaag agggtaaaag aacagttaaag agaaaaatat gataaggagg 240
gggttcagtg agagtgaag ggggtaagag aacagttaaag aaaaaatat gacaaggagg 300
gccatgagga tctacgattc tagttacttt cctcacgggt gtcgcttgaa gacaggtgc 360
agatcctcta gaggttcaca ggaatagcta gcgttgctc ctggattttc gggttccttt 420
ggcagtatac agagtgtgac tcgagtgtga tgtattcaag actccactcc agccacttta 480
accgcagttg gggtagataa aatgactggg taggtgctt cccaggatgt atctaaggat 540
ggggacttag aagggaaggga cttgactaat accatgtcac cagggtgcaa taattacttt 600
ccctcttctc gggaacaggt tccctgtaat gttttaagaa cttgttgata tttggccaag 660
gaggtgatgt ctgcaactaa gctggccatc tctcggtcaa gcacaaggtc cttggttagg 720
aagggccatc catacagcat tttgtatggg ctaagtccct ctttttgggg agagttttg 780
attcttagta aggtgtagg caacagagca ggccatgcaa ggtgggtttc ttgggttagc 840
ttttttaaat gtcgtttgag tgcttcattc attttcttga cttttcttga ggattgtggc 900
ctccacgcgc agtgtaagt atattgtatg cctaagtccct gggatactcc ctgggttact 960
gtagccttga aaacggggcc attgtcactc tgtaagcctc ggggaagtc gaactctggg 1020
attatttcat gaattagtgc ctttattaca tcttggtcct tttctgtcct acaaaggagg 1080
gcctctgccc aaccagtga aatatctacc cagactagta gatactgaaa tccctgagat 1140
ttgggcatgt gggtaaaatc tagttgccag tcttctcctg agtaatggcc tgttctttgt 1200
tctcctgaag gagctt 1216

```

<210> 8  
 <211> 976  
 <212> ADN  
 <213> Homo sapiens

```

<400> 8
agtataatg gaactattga aagtaatccc ctactccag gaactagtgc tgagctggcc 60
aaactaatag ccctcactcg ggcactagaa ttaggagaag agaaaagggt aaatatatat 120
acagactata agtatgctta cctagtcctt catgcccag cagcaatatg gagagaaagg 180
gaattcctaa cttccaaagg aacacctatc aaacatcagg aagccattag gatattatta 240
ttggtgttac agaaacctaa agaggtggca gtcctacact gctgggtgca tcagaaaaaa 300
aaggaaaagg aaatagaagg gaactaccaa gcagatattg aagccaaaag agccgcaagg 360
caggaccctc cattagaaat gcttatagaa ggaccctag tgtgggtgaa cccctccag 420
gaaagcaatc cccagtactc agcaggagaa ataaaatgga gaacctcacg aggacatact 480
ttcctcccct caggatggct agccaccaa gaaggaaaaa tgcttttgcc tgcagctaac 540
caatggaaat tacttaaaac ccttcaccaa acctttcact taggattgat agcaccatc 600

```

```

agatggccaa attattatctt actggatcag gccttttcaa aactatcaag caggtagtca 660
gggcctgtaa agtgtgccaa agaaataatc tcctgcactg caagccatcc atttcaatcc 720
ctgtatcttt aacctccttg ttaagtttgt ctctccaga atcaaaagctg taaaactaca 780
aatgggtctt caaatggagt ctcatatgca gtccatgact aagatatacc gcagccccct 840
ggagggggcc tgctagccca tgctccaatg ttaatgacat cgaaggcacc cctcccgggg 900
aaatctcaac tgcacaaccc ctactatgtc ccaattcagc aggaagcagt taaagcgggtc 960
atcggccaaac ctcccc

```

```

<210> 9
<211> 942
<212> ADN
<213> Homo sapiens

```

```

<400> 9
agaggagaac agcagcataa gcggtctggca gaggtagggga aagaccagca agaagaaaaa 60
agagaaaagag aaagagaaaag tcagagaaaag agacagagag aggaagagac aaagagacag 120
aaagtcaaaag aggtagtagt cagaaacaga gacaaaaaaa agggagtcaga aagaggggaca 180
gacacagaaa gtcaaaaaaaa aagttaagaa gaaaaggaaaa gacaaaagaag aagtcgaaga 240
ggagaaaagag agagatagaa gtagtaagaa aaaaaacagc atatcccatc cctttaaagc 300
cagggtaaat ttctatctac ccagccaagg catattctac ttatgtggat cttcaacca 360
tatctgcctc tcagacagtt tgcaagaaat aatgaaatct atccttactt tacaatccca 420
aatagactct ttggcagcag tgactctcca aaactgcaga ggcctagacc tcctcactgc 480
tgaaaaagga ggacactaca ccttcttagg ggaagaatgt tgtttttaca ctaaccagtc 540
ggggatagta tgagatgctg cccggagttt acaggaaaaag gcttctgaaa tcagacaacg 600
cctttcaaat tcttatacca acttctggag ttaggcaaca tggcttctcc cctttctagg 660
tcctgtggca gccatcttgc tgttactcgc ctttggggccc tgtattttta accttctgtc 720
caaatttgtt tcctctagaa tcgaggccat caagctacag atggtcttac aaatggaacc 780
ccaaaagagt tcaactaaca acttctaccg aggaccctg gatcaaccca ctggcacttc 840
ccctggccta gagagttccc ctctgaagga caccgcaact tgaagggccct tctttgcccc 900
atccagcag agtagctaga gtggtcatcg gccaaattgc ca
942

```

```

<210> 10
<211> 1375
<212> ADN
<213> Homo sapiens

```

```

<400> 10
ccccaatatt ctctttctga tggggaaaaa tggccacctg agggaagcac aaattacaat 60
actatcctgc agcttgatct tttctgtaag agggaaggca aatggagtga aataccttat 120
gtccaagctt tcttttcatt gagggagaat acacaactat gcaaagcttg caattttacat 180
cccacaggag gacccctcag cttaccccca tatcctagcc tccctatagc tcccttctct 240
attgatgata ctctctctct aatctccctt gcccagaagg aaataagcaa agaaatctcc 300
aaagggtccac aaaaaccccc gggctatcgg ttatgtcccc ttcaagctgt agggggaggg 360
gaatttggcc caaccgggtt gcatgtcccc ttctccctct ctgattttaa gcagatcagg 420
cagacctggg gaagttttca gatgatcctg ataggtacat agatgtccta cagggtctag 480
ggcaaacctt tgacctcact tggagagacg tcatgtactt gttagatcaa accctggcct 540
ttaatgaaaa gaatgcggct ttagctgcag cctgagagtt tggagatacc tggatccta 600
gtcaagtaaa tgaaagaatg acagccgaag aaagggacaa cttccctact ggtcagcaag 660
ccatccccag tatggatccc cactgggact ttgactcaga tcatggggac tggagtctga 720
aacatctgtt gatctgtgtt ctggaaggac taaggagaat tgggaaaaag cccatgaatt 780
attcaatgat atccaccata acccaggga aggaagaaaa tccttctgcc ttcctcagac 840
ggctacaaga ggccttaaga aaatatactc ccctgtcacc cgaatcactc gaggttcaat 900
tgattctaaa gaataagttt attaccctaat cagccacaga tatcaggaga aagctccaaa 960
agcaagccct gagccctgaa caaaatctag agacattatt aaacctggca accttggtgt 1020
tctataatag ggaccaagag gaacaggccc aaaaaggaaaa gcgagatcag agaaaggccg 1080
cagccttagt catggccctc agacaaaaca accttggtgt ttcagagagg tcagaaaatg 1140

```

```

gagcaggcca atcacctggt acggcttgtt atcagtgcgg ttactagga cactttaaaa 1200
aagattgtcc aataagaaac aagctgcccc ctcatccgtg tccactatgc cgaggcaatc 1260
actggaaggt gcactgcccc agaggatgaa gggtccctgg gttagaagcc cccaaccaga 1320
tgatccaaca acaggactga ggggtccccg ggcaagcacc agctcatgtc atcac 1375

```

```

<210> 11
<211> 944
<212> ADN
<213> Homo sapiens

```

```

<400> 11
acctaggagg aactgtcttc aggacaggac tatagatgct tcctcccagg cgattaaggg 60
aaaaagacac aatgggtatt cagtaagtga taaggaaact ctgtagaag cagagttagg 120
aaaattgcct aataattggt ctgctcaaat gtgcgagctg ttgcaactca gccaaacctt 180
aaaagtatta cagaatcagg aagaagccat ctataccaat tctaagttaa tatggactga 240
acgagaactt ataatagca aagaataatt gaaatcccaa acttacaagg ttttcaacaa 300
aagcacagtt tgctaaaagt taactgtgta acatgtatta tcctactacc aaaaactctc 360
aaatgatttc tcagacagtt tgcaagaaac aatgaaacct atccttactc tacaatccca 420
aatagactct ttggcagcag tgactctcca aaaccaccaa ggcctagacc tcctactgc 480
tgagaaagga ggactctgca ccttcttagg ggaagattgt tgtttttaca ctaaccagt 540
agggatagtg tgagatgcca cccagcgttt acaggaaaag gcttctgaaa tcagacacaa 600
tgcttttcaa accttatagc aacctctgga gttcggcgac tggcttttcc cctttctagg 660
tcctgtgaca gccatcttgc tattactcgc cttcggggccc tgtattttta acctcctcgt 720
caaatgtgtt tcctctagga tcgaggccat caagctacag atggtcttac aaatggaacc 780
ccaaatgagc tcgactaaca acttctactg aggaccctg gaccgaccca ctggcccttt 840
aactggctta aagagtttcc ctctggagga cactacaact gcaggggccc ttctttgccc 900
catccacagg aagttagcta gagcagtcac cacccaattc ccaa 944

```

```

<210> 12
<211> 963
<212> ADN
<213> Homo sapiens

```

```

<400> 12
tacaggaacc ccataatacg tccttggaac attctattca gctccaactg ctaggagtgg 60
cccatttgtc ctgaaccctc aaatcatggg aatgagaaat gaatttagac tgaccacagc 120
ccttatgagt tttcagctac aggggtgtat agaaccctga taaggagtgt tctttgtgtg 180
tggaagatcc ttctatattt gcctccccac caactggaca ggaacttgta ctttagccta 240
catagtacct cctgtgactt atccttttca gaagaggcag tagctgtgcc cattcatgct 300
aagcttcagc cgagagcaat ctactacttt cctctatttg ctggttttag atttactacc 360
acctaggaag tggactcaca gcctagatga aatctctctc caacttactc aaatccagga 420
ccaaatagac tcattagcag ctgtggttct ccgaaccagt gagcactaga tctccaatct 480
cctcactgcc gaaaggggag gaacatgcct tttctgaac aaggaatgtt gtttttatgt 540
caataaatca ggcatagtga gagatggaat taaatgactt caggatagag ctacgagact 600
acatgggtgg acaaccgaaa ctacctcagg gttctcacag cctgttctcc actggcttct 660
tcatttttta ggtcccttcc ttatgattat tctaggagta acctttggcc catgtctttt 720
cagtcccttc atcctttcgt ttcttctga atagaatcaa tgaaactaga aatgttactg 780
cagatggaac ctcatagcag ttcaaccagc acctattatc aaggaccctt aaaccagcct 840
gccggcccat acccgacgt tgacacccaa accacctctc acgaggaaac ctcatctaca 900
gaacccttcc tatgccccta ttcagcagga agcaattaga gtggtcatcc tcccacaccc 960
caa 963

```

<210> 13  
 <211> 1362  
 <212> ADN  
 <213> Homo sapiens

<400> 13  
 ccacaatatc ctcttccagg aggagaacga tggccacctg aggggaagtat acactataat 60  
 accatcctgc aactagatct gttttgtaaa caagaaggca agtggattta ggtaccatat 120  
 gttcagacct ttttctcatt aagggatgat aaccacgat tgtgtaagac atgtaacctg 180  
 caccacacag ggagtcctca aattctaccc ccataccag tcctcccccac ggctcctcct 240  
 actaatgcca aacctctctt ggcttctaca gcccaaaagg gaacaaataa aagagccttc 300  
 agagagccaa gagacccac tggcccctgg ctatgtcctc ttcaggctgt aggggggaa 360  
 tttggcccaa ccgagtaga tgttcccttt tctctctctg atctaaagca aattaaggca 420  
 gacttggatg aaagtctca gatgaccca atagatacgt agatggcctg ctgggtctgg 480  
 gacaatcttt tgacctttcc tggagagaga tcatgttatt gcttgatcag acctaacctc 540  
 taatgagaag aatgctgctt taacaggagc ccgagagttt ggggatacct ggtacctcag 600  
 ttaagtaagt gatagaatga catcagaaga gagcagtttc ctactggcca gcaagcagtc 660  
 cccagtatgg atccccactg ggaccctgac tcggatcatg gggactggag tcacaaaact 720  
 ttactgacct gtatcctaga agggtttaagg agaactagga aaaaagccat gaactattca 780  
 atgatgtcta ctataaccac aggggaaggaa gaaaacccta ttgccttcct caaaaggctg 840  
 agggaggctt tgagaaaata tactccctcg tcaccagatt ccctcgaagg ccagttaatt 900  
 ttaaaggaca aatttattac tcagtcagct gcagacatta ggaaaaagct ccaaaagtta 960  
 gccttgggcc gagcaaaatt tggaggcatc attaaacctg gcaacctcag tgttctatca 1020  
 tagggaccaaa gaggaacagg ccgaaaagga aaagcaggat aagagaaagg ctgcagattt 1080  
 agtcatgccc tcagacaaac cttggcggtt caaagaggag aaaaaatgga gcaggccaat 1140  
 caccacgacg ggcttattat cagtgcagtt tacaaggaca ctttaaaca gattgtccaa 1200  
 agagaaataa gccgccctct caccatgtc cactatgcca agtgatcac tggaaaggac 1260  
 actgtccag aggacaaaagg ttctctgggc cagaagtccc caaccagatg atccagcaac 1320  
 aggatggagg gtgcccggg caagcaccag ctctgttgtt ca 1362

<210> 14  
 <211> 945  
 <212> ADN  
 <213> Homo sapiens

<400> 14  
 ttgcagatca atctcagact gctgtgctag caatgagtga ggcttcgtgg gcatgggacc 60  
 ctctgagcca ggcattggat ataatgtcct tgtgtgccat ttgctaagac tgttggaata 120  
 gcacagtatt aggggtggag tggcccatt ttccagggtc tgtctgtcac cgcttccctt 180  
 ggctagggaa gagaattccc tgacctcttg ttcttcccag gtaaggcagt gcctcaccct 240  
 gcttcagctc acactcaggt gactgcaccc actgtcctgc cccactgtc ggacaagccc 300  
 cagttagatg aacctggtac ctacgttga aatgcagaaa tcacctgtct tctgcgtcac 360  
 tcacactggg agctgtagac tggagctgtt cctatttggc catcttgga ccactctcca 420  
 aatagactct ttggcagcag tgactctcca aaaccacca ggcttagacc tcctcattgc 480  
 tgagaaagga ggactctgca ccttcttagg ggaggagtgt tgtttttata ctgaccagtc 540  
 agggatggta cgagatgcca cccgatgtt acaggaaaag gcttctgaaa tcacacaaca 600  
 cctttcaaac tcttatacca acctctggag ttgggcaaca tggcttctcc cctttctcgg 660  
 tcccattgca gccatcttgc tattactcgc cttcaggctg tgtattttta acctccttgt 720  
 caaatttgtt tcctctagaa ttgaggccgt caagctacag atggctttac aaatgggacc 780  
 ccaaatgagc tcaactaaca acttctgcca aggacccctg gaccaacctg ctggcccttt 840  
 cactggcctt aagagtccc ctctggaggg cactacaact gcagggcccc ttctttgccc 900  
 ctatccagca ggaagtagct agagcagtca tcaccaatt cccaa 945

<210> 15  
 <211> 939  
 <212> ADN  
 <213> Homo sapiens

<400> 15  
 agagctacct tggcaagtac tctaggagta tgggaaaatg aaaacaacaa actcacacac 60  
 cattttaaca tacacaatca ggtctgcca cccagcaagg tatattcttt gtatgtggaa 120  
 catcgacctt tatctgcctc cccactaact agacagccac ctgaatctta gtctttctaa 180  
 gtcccaacag taacattgcc ccaggaaatc agaccatata agtatccctc aaagctcaa 240  
 tctgtcagtg cagagccata caactaatc ccctacttat agggtaagga atggctact 300  
 ctacaggaac cagaatagct agtttggtta ctctattatc ctactaccac acactctcaa 360  
 atgatttctc agacagtgtg caagaaataa cgaaatctat ccttactcta caatcccaaa 420  
 tagactcctt ggcagcagtg accctccaaa acggctgagg cctagacctc ctactgcca 480  
 agaaaaggagg actctgcatt ttcttagggg aagagtgttt ttacactaac cagtccaggga 540  
 cagtatgaga tgccactcgg agtttacagg aaaaggcttc tgaagtcaga caatgcct 600  
 caaactctat accaaactct ggagttgggc aacatggctt ctccctcttc taggtccc 660  
 gacagccatc ttgctattat ttgcctttga gccctgtatt tttaatctcc ttttcaaa 720  
 tgtttcctct ggatcgaggc catcgagcta cagatggtct tcacaaatgg aaccccaaa 780  
 gagctcaact aacaacttct actgaggacc cctggactaa cctgctgacc ctttcaactg 840  
 cctgaagaat tcccctctgg aggacactac aactgcaggg ctccctcttt gccctatcc 900  
 agcaggaagt agctagagct gtcattgcct aattcctaa 939

<210> 16  
 <211> 979  
 <212> ADN  
 <213> Homo sapiens

<400> 16  
 agtgataatg gaatacttga aagtaatccc ctactcccc aggaactagt gctcagctgg 60  
 cagaactaat agccctcact cgggtactag aatcaggaga aggaaaaagg gtaaatatat 120  
 atacagactc taagtgtgct tacctagtcc tccatgccca tgcagcaata tggagagaaa 180  
 gggaattcct aacttccgag ggaacaccta tcaaaccatca ggaagccatt aggaaattat 240  
 tattggctgt acagaaacct aaagaggtgg cagttttaca ctgccggggg catcagaaa 300  
 gaaaggaaa ggaaatacaa gggagccacc aagttgatat tgaagtcaaa agagccacaa 360  
 ggctggaccc tccattagaa atgcttatag gaggaccctc agtatggggg aatccctcc 420  
 gggaagccaa gccccagtag tcagcaggag aaatagaata gggaaacttc tgaggacata 480  
 ctccctccc ctccagatgg ctagccacca ataaaggaaa aatacttttg cctgcagcta 540  
 accaatagaa attacttaaa acccttcac aaacctcca cttaggcatt gatagcacc 600  
 atgagatggc caaattatta ttactggac caggcctttt caaaactatc aagcagatag 660  
 tcagggcctg taaagtctgc caaagaaata atcccctgca ctgcaggcca tacatttcaa 720  
 tccctgtatc ttttaacctc ttcttaaat tgtctctcc agaatacaa ctgtaaaatt 780  
 acaaatagtt cttcaaatgg agccacagat gcagtccatg actaagatcc accacagacc 840  
 cctggaccag cctgctagcc catgctccaa tgtaatgac atcgaaggca cccctcctg 900  
 aggaaatctc aactgcacaa cccctactac gcccgaattc agcagaaagc agttagagt 960  
 gtcacagcc aacctcccc 979

<210> 17  
 <211> 1774  
 <212> ADN  
 <213> Homo sapiens

<400> 17  
 catgctggta aaggaccgct agaatccagc agccaggacc actttctttg tggtaagaa 60  
 aggtgggaaa acaggtgcag gactgtaca ctggtgaagca taactaatcc gataagcaga 120  
 ggccatggg tggttacgca ccctggaaag gaataagcat taggactata gaggacactc 180

```

taggactaat gctcatcgga aaatgactag ggggtactggc atccctatgc tcttttttca 240
gatgggaaat gttccccca aggcagaaat gccctaaga tgtattctcg agaaatggga 300
ccaatctgac catcagacac taagaaagaa atgacttata tcttctcgca gtaccacctg 360
gccacaatat cttcttcaag gggcagaaac ctggcctcct gaggggaagta taaattataa 420
caccatctta cagctagacc tcttttgtag aaaagaaggc aaatggagtg aagtgccata 480
tgtatacaat tcttttcat taagagataa ctccaatta tgtaaaaagt gtgatttatg 540
ccctacagga agccctcaga gtctacctcc cgaccccagc aagaccccaa ctccttctcc 600
aactaataag gacccccctt caacccaaat ggtccaaaag gagatagaca aaggggtaaa 660
caatgaacca aagagtgcc aatattacacg attatactcg ctccaagcag tgggaggaga 720
atthggccca gccagcgtgc atgtaccttt tctctctca gatttaaagc aaattaaaaa 780
agacctaggt aaattctcag ataaccctga tggctatatt gatgttttac aagggttagg 840
acaatccttt gatctgacat ggagagatat aatgttactg ctaaatcaga cactaaccctc 900
aaatgaaaaa agtgctgcc taacagcagc ctgagagttt ggcgaactct ggtatctcag 960
tcaggccaat gataggatga caacagatga aagagaatga ttccccacag gccagcaggc 1020
agttcccagt gtagaccctc attaggacac agaatacaga cttggagatg ggtgccacag 1080
acatttgcta acttgctgct tagaaggact aaggaaaact aggaagaagc ccatgaatta 1140
ttcaatgatg tcccctataa cacagggaaa ggaagaaaat cctactgcc tcttgagag 1200
actaaggaaa ggattgagga agcatacctc cctgtcacct gactctatta aaggccaact 1260
aatctttaaag gataagtta tcaactagtc agctgcagag attaagaaaa aacttcaaaa 1320
gtatgcctta ggcccagagc aaaacttaga aaccctactg aacttggaac cctcagtttt 1380
ttataataga gatcaggaag agcaggggaa tgggacaaat gggataaaaa aaaaaaaaaa 1440
aggtgactgc tttagtctgt gccctcaggc aaatggactt tggaggctcc agaaaaggga 1500
aaagctgagc aaattgaatg cctaacaggc cttgcttcta gtgtggtcta caaggacact 1560
ttaaaaaaga ttgtccaagt agaaacaagc tgcccccttg tccatgcccc ttatgtcaag 1620
ggaatcactg gaaggccac tgccccagga gatgaaggtc ctctgagtca gaagccacta 1680
accagataat ccagcagcag gactgaggat gccaggggca agcgccagcc catgccatca 1740
ccctcacaga gccttgggta tgcttgacca ttga 1774

```

<210> 18  
 <211> 938  
 <212> ADN  
 <213> Homo sapiens

```

<400> 18
tgttaggaaga actcccttca ggacaggaca atagatgggt cctcccaggc gattaaggaa 60
aaaagacaca gtattcagta agtgataagg aaactcttgt agaagcagag tttagaaaaa 120
tgcctaataa ttggtctgct caaatgtgtg agttgtttgc actcagccaa atcttaaaat 180
acttacagaa tcaggaagca gccatctata ccaattctaa gttaatatgg actaaacgag 240
gttttattag tagcaaagaa aaattaaaaa cccaaactta caagggtttc aactaaagtt 300
tgccaaaagt taacagtgt aatgttatta tctactatc acacactctc aaaggatttc 360
tcagacagtt tgcaagaaat aacgtaatct atccttactc tacagtccca aatagactct 420
ttggtagcag tgaactctca aaactgccga ggtctagacc tcctcaatgc tgagaaagga 480
gaactctgca cttcttagg ggaagagtgc tgtttttaca ctaaccagtc agggatagta 540
tgagatactg cctgacgttt acaggaaaag gcttctgaaa tcagacaacg ctttcaagc 600
tcttatacca acctctggag ttgggcaaca tggcttctcc ctttgctagg tctgtggca 660
gccatcttgc tattacttgc cttcgggccc tgtattttta acctccttgt caaatttgtt 720
tcctctagga tcaaggccat caagctacag atggtcttac aaatggaacc ccaaatgagc 780
tcaactaaca acttctactg aggcacactg gactgaccca ctggcccttt cactggccta 840
aagagttccc ttctggagga cactacaact gcagggcccc gtcttcaccc ctatccagca 900
ggaagtagct agatcagtca ttgcccatt cccaacag 938

```

<210> 19  
 <211> 1308  
 <212> ADN  
 <213> Homo sapiens

&lt;400&gt; 19

```

gatgcttgcc ccaggcaccc tcagtcctgt tgttggatca tctggctcggg ggcttctggc 60
ccaaagaacc ttgtcctct gaggcagtgc acctccagt gattgcctca gcattgtgga 120
catgggcaag ggggcagcct gtttctcact ggacaatctt ttttaagggtg tccttccaaa 180
ccacactggg aacaagccct accaggtgat tggcctgctc tattttctgt cctctctgaa 240
ccaccaagggt ttgtctgtct gagggcatg actaaggctg tggcctttct ctgatcttgc 300
ttttcctttt tggcctgttc ctcttggtag ctattataga aactgaggt tggcaggtt 360
aacaatggct ccagattttg ttcagggcac agggctcatt ttggagcttt ctctgatata 420
ctgcagctga ttgggtaata aacttatctt ttaggatcaa ttgactctca agagagttgg 480
gtgacaggaat agtatatttc cttgaggcct cccatagccg ctctagggaag gcagaaggat 540
tttcttcctt tccctgagtt ataaaagaca tcattgaaca actcatggac tttttcccaa 600
ttctccgtag tccttctaga acacagggtca gcagatgttt acgactccag tccccatgat 660
ctgagcttag acaccagtgg ggatccatac tggggatggc ctgctgactg gtaggggaatt 720
tgtccctttc ttggctgtc attctatcat ttacttgact aagataccaa gtatctccaa 780
attctcaggt tgcagctaaa gctgcattct tttcattaaa ggccagggtt tgatctaata 840
gcatgacatc tctccaagtg aggtcaaagg tttgccctag atccatagga catcagagaa 900
ggagaagggg acatacacct gagttagcca aattcccctc cctctacagc ttgaagggga 960
cataagcaat agcctgggga tttttgtgtt cctttggaga tttctttgct tgtttccttc 1020
tgggtggggg agattagagg aggccttatca gtaataggaa ggggagctat agggaggcta 1080
ggatatgggg gtaagctgag aggtcatctt gtgggatgta aattgcaagc ttgcatagt 1140
tgtggatttt ccttacaatg aaaataaagc ttggacataa ggtatttcac tccatttgcc 1200
ttccctctta cagaaaagggt caagctgcag gatagtactg taatttatac ttccttcagg 1260
tggccatttc ttcccatcag agagagaata ctggggctgg gccatagt 1308

```

&lt;210&gt; 20

&lt;211&gt; 711

&lt;212&gt; ADN

&lt;213&gt; Homo sapiens

&lt;400&gt; 20

```

actgagagac aggactagct ggatttccta ggccgactaa gaatccctaa gcctagctgg 60
gaaggtgacc acgtccacct ttaaaccacgg ggcttgcaac ttagctcaca cctgaccaat 120
cagagagctc actaaaatgc taattaggca aagacaggag gtaaaagaaat agccaatcat 180
ctattgcctg agagcacagc aggaggggaca acaatcgga tataaaccba ggcattcgag 240
ctggcaacag cagccccctt ttgggtccct tccctttgta tgggagctgt tttcatgcta 300
tttactctta ttaaactctg caactgcact ctctgggtcc atgtttctta cggctcgagc 360
tgagcttttg ctcaccgtcc accactgctg tttgccacca ccgcagacct gccgctgact 420
cccatccctc tggatcctgc aggggtgtccg ctgtgctcct gatccagcga ggcgccatt 480
gccgctccca attgggctaa aggccttgcca ttgttcctgc acggctaagt gcctgggttt 540
gttctaattg agctgaacac tagtcactgg gttccatggt tctcttctgt gaccacggc 600
ttctaataga actataacac ttaccacatg gcccaagatt ccattccttg gaatccgtga 660
ggccaagaac tccagggtcag agaatacgag gcttgccacc atcttggaag c 711

```

&lt;210&gt; 21

&lt;211&gt; 711

&lt;212&gt; ADN

&lt;213&gt; Homo sapiens

&lt;400&gt; 21

```

actgagagac aggactagct ggatttccta ggctgactaa gaatccctaa gcctagctgg 60
gaaggtgacc acatccacct ttaaaccacgg ggcttgcaac ttagctcaca cctgaccaat 120
cagagagctc actaaaatgc taattaggca aagacaggag gtaaaagaaat agccaatcat 180
ctattgcctg agagcacagc aggaggggaca atgatcgga tataaaccba agtcttcgag 240
ccggcaacgg caacccccct tgggtccctt cctttgtat gggagctctg tttcatgct 300
atttactctt attaaactct gcaactgcac tcttctggct catgtttctt acggcttgag 360
ctgagctttc gctcgccatc caccactgct gtttgccgcc accgcagacc cgccgctgac 420

```

15

```

tcccatccct ctggatcatg caggggtgtcc gctgtgtctcc tgateccagcg aggcacccat 480
tgccgctccc aatcgggcta aaggcttgcc attgttcctg catggctaag tgcctgggtt 540
catcctaatt gagctgaaca ctatgactg ggttccatgg ttctcttctg tgacccacag 600
cttctaatag agctataaca ctacccgcat ggccaaggt tccattcctt gaatccataa 660
ggccaagaac cccaggtcag agaacacgag gcttgccacc atcttgggag c 711

```

<210> 22  
 <211> 2055  
 <212> ADN  
 <213> Homo sapiens

<220>  
 <221> CDS  
 <222> (1)..(2055)

<400> 22  
 ccc aag aca gcc aac tta gtt gca gac atc acc tcc tta gcc aaa tat 48  
 Pro Lys Thr Ala Asn Leu Val Ala Asp Ile Thr Ser Leu Ala Lys Tyr  
 1 5 10 15  
 caa caa gtt ctt aaa aca tta caa gga acc tat ccc tga gaa gag gga 96  
 Gln Gln Val Leu Lys Thr Leu Gln Gly Thr Tyr Pro Glu Glu Gly  
 20 25 30  
 aaa gaa cta ttc cac cct tgt gac atg gta tta gtc aag tcc ctt ccc 144  
 Lys Glu Leu Phe His Pro Cys Asp Met Val Leu Val Lys Ser Leu Pro  
 35 40 45  
 tct aat tcc cca tcc cta gat aca tcc tgg gaa gga ccc tac cca gtc 192  
 Ser Asn Ser Pro Ser Leu Asp Thr Ser Trp Glu Gly Pro Tyr Pro Val  
 50 55 60  
 att tta tct acc cca act gcg gtt aaa gtg gct gga gtg gag tct tgg 240  
 Ile Leu Ser Thr Pro Thr Ala Val Lys Val Ala Gly Val Glu Ser Trp  
 65 70 75 80  
 ata cat cac act tga gtc aaa tcc tgg ata ctg cca aag gaa cct gaa 288  
 Ile His His Thr Val Lys Ser Trp Ile Leu Pro Lys Glu Pro Glu  
 85 90 95  
 aat cca gga gac aac gct agc tat tcc tgt gaa cct cta gag gat ttg 336  
 Asn Pro Gly Asp Asn Ala Ser Tyr Ser Cys Glu Pro Leu Glu Asp Leu  
 100 105 110  
 cgc ctg ctc ttc aaa caa caa cca gga gga aag taa cta aaa tca taa 384  
 Arg Leu Leu Phe Lys Gln Gln Pro Gly Gly Lys Leu Lys Ser  
 115 120 125  
 atc ccc atg gcc ctc cct tat cat att ttt ctc ttt act gtt ctt tta 432  
 Ile Pro Met Ala Leu Pro Tyr His Ile Phe Leu Phe Thr Val Leu Leu  
 130 135 140  
 ccc tct ttc act ctc act gca ccc cct cca tgc cgc tgt atg acc agt 480  
 Pro Ser Phe Thr Leu Thr Ala Pro Pro Pro Cys Arg Cys Met Thr Ser  
 145 150 155 160



16

agc tcc cct tac caa gag ttt cta tgg aga atg cag cgt ccc gga aat	528
Ser Ser Pro Tyr Gln Glu Phe Leu Trp Arg Met Gln Arg Pro Gly Asn	
165 170 175	
att gat gcc cca tcg tat agg agt ctt tct aag gga acc ccc acc ttc	576
Ile Asp Ala Pro Ser Tyr Arg Ser Leu Ser Lys Gly Thr Pro Thr Phe	
180 185 190	
act gcc cac acc cat atg ccc cgc aac tgc tat cac tct gcc act ctt	624
Thr Ala His Thr His Met Pro Arg Asn Cys Tyr His Ser Ala Thr Leu	
195 200 205	
tgc atg cat gca aat act cat tat tgg aca gga aaa atg att aat cct	672
Cys Met His Ala Asn Thr His Tyr Trp Thr Gly Lys Met Ile Asn Pro	
210 215 220	
agt tgt cct gga gga ctt gga gtc act gtc tgt tgg act tac ttc acc	720
Ser Cys Pro Gly Gly Leu Gly Val Thr Val Cys Trp Thr Tyr Phe Thr	
225 230 235 240	
caa act ggt atg tct gat ggg ggt gga gtt caa gat cag gca aga gaa	768
Gln Thr Gly Met Ser Asp Gly Gly Gly Val Gln Asp Gln Ala Arg Glu	
245 250 255	
aaa cat gta aaa gaa gta atc tcc caa ctc acc cgg gta cat ggc acc	816
Lys His Val Lys Glu Val Ile Ser Gln Leu Thr Arg Val His Gly Thr	
260 265 270	
tct agc ccc tac aaa gga cta gat ctc tca aaa cta cat gaa acc ctc	864
Ser Ser Pro Tyr Lys Gly Leu Asp Leu Ser Lys Leu His Glu Thr Leu	
275 280 285	
cgt acc cat act cgc ctg gta agc cta ttt aat acc acc ctc act ggg	912
Arg Thr His Thr Arg Leu Val Ser Leu Phe Asn Thr Thr Leu Thr Gly	
290 295 300	
ctc cat gag gtc tcg gcc caa aac cct act aac tgt tgg ata tgc ctc	960
Leu His Glu Val Ser Ala Gln Asn Pro Thr Asn Cys Trp Ile Cys Leu	
305 310 315 320	
ccc ctg aac ttc agg cca tat gtt tca atc cct gta cct gaa caa tgg	1008
Pro Leu Asn Phe Arg Pro Tyr Val Ser Ile Pro Val Pro Glu Gln Trp	
325 330 335	
aac aac ttc agc aca gaa ata aac acc act tcc gtt tta gta gga cct	1056
Asn Asn Phe Ser Thr Glu Ile Asn Thr Thr Ser Val Leu Val Gly Pro	
340 345 350	
ctt gtt tcc aat ctg gaa ata acc cat acc tca aac ctc acc tgt gta	1104
Leu Val Ser Asn Leu Glu Ile Thr His Thr Ser Asn Leu Thr Cys Val	
355 360 365	
aaa ttt agc aat act aca tac aca acc aac tcc caa tgc atc agg tgg	1152
Lys Phe Ser Asn Thr Thr Thr Thr Asn Ser Gln Cys Ile Arg Trp	
370 375 380	

gta act cct ccc aca caa ata gtc tgc cta ccc tca gga ata ttt ttt	1200
Val Thr Pro Pro Thr Gln Ile Val Cys Leu Pro Ser Gly Ile Phe Phe	
385 390 395 400	
gtc tgt ggt acc tca gcc tat cgt tgt ttg aat ggc tct tca gaa tct	1248
Val Cys Gly Thr Ser Ala Tyr Arg Cys Leu Asn Gly Ser Ser Glu Ser	
405 410 415	
atg tgc ttc ctc tca ttc tta gtg ccc cct atg acc atc tac act gaa	1296
Met Cys Phe Leu Ser Phe Leu Val Pro Pro Met Thr Ile Tyr Thr Glu	
420 425 430	
caa gat tta tac agt tat gtc ata tct aag ccc cgc aac aaa aga gta	1344
Gln Asp Leu Tyr Ser Tyr Val Ile Ser Lys Pro Arg Asn Lys Arg Val	
435 440 445	
ccc att ctt cct ttt gtt ata gga gca gga gtg cta ggt gca cta ggt	1392
Pro Ile Leu Pro Phe Val Ile Gly Ala Gly Val Leu Gly Ala Leu Gly	
450 455 460	
act ggc att ggc ggt atc aca acc tct act cag ttc tac tac aaa cta	1440
Thr Gly Ile Gly Gly Ile Thr Thr Ser Thr Gln Phe Tyr Tyr Lys Leu	
465 470 475 480	
tct caa gaa cta aat ggg gac atg gaa cgg gtc gcc gac tcc ctg gtc	1488
Ser Gln Glu Leu Asn Gly Asp Met Glu Arg Val Ala Asp Ser Leu Val	
485 490 495	
acc ttg caa gat caa ctt aac tcc cta gca gca gta gtc ctt caa aat	1536
Thr Leu Gln Asp Gln Leu Asn Ser Leu Ala Ala Val Val Leu Gln Asn	
500 505 510	
cga aga gct tta gac ttg cta acc gct gaa aga ggg gga acc tgt tta	1584
Arg Arg Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly Thr Cys Leu	
515 520 525	
ttt tta ggg gaa gaa tgc tgt tat tat gtt aat caa tcc gga atc gtc	1632
Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val	
530 535 540	
act gag aaa gtt aaa gaa att cga gat cga ata caa cgt aga gca gag	1680
Thr Glu Lys Val Lys Glu Ile Arg Asp Arg Ile Gln Arg Arg Ala Glu	
545 550 555 560	
gag ctt cga aac act gga ccc tgg ggc ctc ctc agc caa tgg atg ccc	1728
Glu Leu Arg Asn Thr Gly Pro Trp Gly Leu Leu Ser Gln Trp Met Pro	
565 570 575	
tgg att ctc ccc ttc tta gga cct cta gca gct ata ata ttg cta ctc	1776
Trp Ile Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile Ile Leu Leu Leu	
580 585 590	
ctc ttt gga ccc tgt atc ttt aac ctc ctt gtt aac ttt gtc tct tcc	1824
Leu Phe Gly Pro Cys Ile Phe Asn Leu Leu Val Asn Phe Val Ser Ser	
595 600 605	

18

```

aga atc gaa gct gta aaa cta caa atg gag ccc aag atg cag tcc aag 1872
Arg Ile Glu Ala Val Lys Leu Gln Met Glu Pro Lys Met Gln Ser Lys
    610                615                620

act aag atc tac cgc aga ccc ctg gac cgg cct gct agc cca cga tct 1920
Thr Lys Ile Tyr Arg Arg Pro Leu Asp Arg Pro Ala Ser Pro Arg Ser
    625                630                635                640

gat gtt aat gac atc aaa ggc acc cct cct gag gaa atc tca gct gca 1968
Asp Val Asn Asp Ile Lys Gly Thr Pro Pro Glu Glu Ile Ser Ala Ala
    645                650                655

caa cct cta cta cgc ccc aat tca gca gga agc agt tag agc ggt cgt 2016
Gln Pro Leu Leu Arg Pro Asn Ser Ala Gly Ser Ser Ser Gly Arg
    660                665                670

cgg cca acc tcc cca aca gca ctt agg ttt tcc tgt tga 2055
Arg Pro Thr Ser Pro Thr Ala Leu Arg Phe Ser Cys
    675                680                685

```

<210> 23  
 <211> 28  
 <212> PRT  
 <213> Homo sapiens

<400> 23  
 Pro Lys Thr Ala Asn Leu Val Ala Asp Ile Thr Ser Leu Ala Lys Tyr  
     1                    5                    10                    15  
 Gln Gln Val Leu Lys Thr Leu Gln Gly Thr Tyr Pro  
             20                    25

<210> 24  
 <211> 55  
 <212> PRT  
 <213> Homo sapiens

<400> 24  
 Glu Glu Gly Lys Glu Leu Phe His Pro Cys Asp Met Val Leu Val Lys  
     1                    5                    10                    15  
 Ser Leu Pro Ser Asn Ser Pro Ser Leu Asp Thr Ser Trp Glu Gly Pro  
             20                    25                    30  
 Tyr Pro Val Ile Leu Ser Thr Pro Thr Ala Val Lys Val Ala Gly Val  
             35                    40                    45  
 Glu Ser Trp Ile His His Thr  
     50                    55

<210> 25  
 <211> 38  
 <212> PRT  
 <213> Homo sapiens

19

&lt;400&gt; 25

Val Lys Ser Trp Ile Leu Pro Lys Glu Pro Glu Asn Pro Gly Asp Asn  
 1 5 10 15

Ala Ser Tyr Ser Cys Glu Pro Leu Glu Asp Leu Arg Leu Leu Phe Lys  
 20 25 30

Gln Gln Pro Gly Gly Lys  
 35

&lt;210&gt; 26

&lt;211&gt; 540

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 26

Ile Pro Met Ala Leu Pro Tyr His Ile Phe Leu Phe Thr Val Leu Leu  
 1 5 10 15

Pro Ser Phe Thr Leu Thr Ala Pro Pro Pro Cys Arg Cys Met Thr Ser  
 20 25 30

Ser Ser Pro Tyr Gln Glu Phe Leu Trp Arg Met Gln Arg Pro Gly Asn  
 35 40 45

Ile Asp Ala Pro Ser Tyr Arg Ser Leu Ser Lys Gly Thr Pro Thr Phe  
 50 55 60

Thr Ala His Thr His Met Pro Arg Asn Cys Tyr His Ser Ala Thr Leu  
 65 70 75 80

Cys Met His Ala Asn Thr His Tyr Trp Thr Gly Lys Met Ile Asn Pro  
 85 90 95

Ser Cys Pro Gly Gly Leu Gly Val Thr Val Cys Trp Thr Tyr Phe Thr  
 100 105 110

Gln Thr Gly Met Ser Asp Gly Gly Gly Val Gln Asp Gln Ala Arg Glu  
 115 120 125

Lys His Val Lys Glu Val Ile Ser Gln Leu Thr Arg Val His Gly Thr  
 130 135 140

Ser Ser Pro Tyr Lys Gly Leu Asp Leu Ser Lys Leu His Glu Thr Leu  
 145 150 155 160

Arg Thr His Thr Arg Leu Val Ser Leu Phe Asn Thr Thr Leu Thr Gly  
 165 170 175

Leu His Glu Val Ser Ala Gln Asn Pro Thr Asn Cys Trp Ile Cys Leu  
 180 185 190

Pro Leu Asn Phe Arg Pro Tyr Val Ser Ile Pro Val Pro Glu Gln Trp  
 195 200 205

Asn Asn Phe Ser Thr Glu Ile Asn Thr Thr Ser Val Leu Val Gly Pro  
 210 215 220

Leu Val Ser Asn Leu Glu Ile Thr His Thr Ser Asn Leu Thr Cys Val  
 225 230 235 240  
 Lys Phe Ser Asn Thr Thr Tyr Thr Thr Asn Ser Gln Cys Ile Arg Trp  
 245 250 255  
 Val Thr Pro Pro Thr Gln Ile Val Cys Leu Pro Ser Gly Ile Phe Phe  
 260 265 270  
 Val Cys Gly Thr Ser Ala Tyr Arg Cys Leu Asn Gly Ser Ser Glu Ser  
 275 280 285  
 Met Cys Phe Leu Ser Phe Leu Val Pro Pro Met Thr Ile Tyr Thr Glu  
 290 295 300  
 Gln Asp Leu Tyr Ser Tyr Val Ile Ser Lys Pro Arg Asn Lys Arg Val  
 305 310 315 320  
 Pro Ile Leu Pro Phe Val Ile Gly Ala Gly Val Leu Gly Ala Leu Gly  
 325 330 335  
 Thr Gly Ile Gly Gly Ile Thr Thr Ser Thr Gln Phe Tyr Tyr Lys Leu  
 340 345 350  
 Ser Gln Glu Leu Asn Gly Asp Met Glu Arg Val Ala Asp Ser Leu Val  
 355 360 365  
 Thr Leu Gln Asp Gln Leu Asn Ser Leu Ala Ala Val Val Leu Gln Asn  
 370 375 380  
 Arg Arg Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly Thr Cys Leu  
 385 390 395 400  
 Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val  
 405 410 415  
 Thr Glu Lys Val Lys Glu Ile Arg Asp Arg Ile Gln Arg Arg Ala Glu  
 420 425 430  
 Glu Leu Arg Asn Thr Gly Pro Trp Gly Leu Leu Ser Gln Trp Met Pro  
 435 440 445  
 Trp Ile Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile Ile Leu Leu Leu  
 450 455 460  
 Leu Phe Gly Pro Cys Ile Phe Asn Leu Leu Val Asn Phe Val Ser Ser  
 465 470 475 480  
 Arg Ile Glu Ala Val Lys Leu Gln Met Glu Pro Lys Met Gln Ser Lys  
 485 490 495  
 Thr Lys Ile Tyr Arg Arg Pro Leu Asp Arg Pro Ala Ser Pro Arg Ser  
 500 505 510  
 Asp Val Asn Asp Ile Lys Gly Thr Pro Pro Glu Glu Ile Ser Ala Ala  
 515 520 525

21

Gln Pro Leu Leu Arg Pro Asn Ser Ala Gly Ser Ser  
 530 535 540

<210> 27  
 <211> 15  
 <212> PRT  
 <213> Homo sapiens

<400> 27  
 Ser Gly Arg Arg Pro Thr Ser Pro Thr Ala Leu Arg Phe Ser Cys  
 1 5 10 15

<210> 28  
 <211> 1080  
 <212> ADN  
 <213> Homo sapiens

<220>  
 <221> CDS  
 <222> (1)..(1080)

<400> 28  
 acc tct ttt gta gaa aag gca aat gga gtg aag tgc cat aag tac aaa 48  
 Thr Ser Phe Val Glu Lys Ala Asn Gly Val Lys Cys His Lys Tyr Lys  
 1 5 10 15

ctt tct ttt cat taa gag aca act cac aat tat gta aaa agt gtg att 96  
 Leu Ser Phe His Glu Thr Thr His Asn Tyr Val Lys Ser Val Ile  
 20 25 30

tat gcc cta cag gaa gcc ttc aga gtc tac ctc cct atc cca gca tcc 144  
 Tyr Ala Leu Gln Glu Ala Phe Arg Val Tyr Leu Pro Ile Pro Ala Ser  
 35 40 45

ccg act cct tcc cca act aat aag gac ccc cct tca acc caa atg gtc 192  
 Pro Thr Pro Ser Pro Thr Asn Lys Asp Pro Pro Ser Thr Gln Met Val  
 50 55 60

caa aag gag ata gac aaa agg gta aac agt gaa cca aag agt gcc aat 240  
 Gln Lys Glu Ile Asp Lys Arg Val Asn Ser Glu Pro Lys Ser Ala Asn  
 65 70 75 80

att ccc caa tta tga ccc ctc caa gca gtg gga gga aga gaa ttc ggc 288  
 Ile Pro Gln Leu Pro Leu Gln Ala Val Gly Gly Arg Glu Phe Gly  
 85 90 95

cca gcc aga gtg cat gtg cct ttt tct ctc cca gac tta aag caa ata 336  
 Pro Ala Arg Val His Val Pro Phe Ser Leu Pro Asp Leu Lys Gln Ile  
 100 105 110

aaa aca gac tta ggt aaa ttc tca gat aac cct gat ggc tat att gat 384  
 Lys Thr Asp Leu Gly Lys Phe Ser Asp Asn Pro Asp Gly Tyr Ile Asp  
 115 120 125

ggt tta caa ggg tta gga caa ttc ttt gat ctg aca tgg aga gat ata	432
Val Leu Gln Gly Leu Gly Gln Phe Phe Asp Leu Thr Trp Arg Asp Ile	
130 135 140	
atg tca ctg cta aat cag aca cta acc cca aat gag aga agt gcc acc	480
Met Ser Leu Leu Asn Gln Thr Leu Thr Pro Asn Glu Arg Ser Ala Thr	
145 150 155 160	
ata act gca gcc tga gag ttt ggc gat ctc tgg tat ctc agt cag gtc	528
Ile Thr Ala Ala Glu Phe Gly Asp Leu Trp Tyr Leu Ser Gln Val	
165 170 175	
aat gat agg atg aca aca gag gaa aga gaa tga ttc ccc aca gcc cag	576
Asn Asp Arg Met Thr Thr Glu Glu Arg Glu Phe Pro Thr Gly Gln	
180 185 190	
cag gca gtt ccc agt cta gac cct cat tgg gac aca gaa tca gaa cat	624
Gln Ala Val Pro Ser Leu Asp Pro His Trp Asp Thr Glu Ser Glu His	
195 200 205	
gga gat tgg tgc tgc aga cat ttg cta act tgt gtg cta gaa gga cta	672
Gly Asp Trp Cys Cys Arg His Leu Leu Thr Cys Val Leu Glu Gly Leu	
210 215 220	
agg aaa act agg aag aag tct atg aat tac tca atg atg tcc acc ata	720
Arg Lys Thr Arg Lys Lys Ser Met Asn Tyr Ser Met Met Ser Thr Ile	
225 230 235 240	
aca cag gga agg gaa gaa aat cct act gcc ttt ctg gag aga cta agg	768
Thr Gln Gly Arg Glu Glu Asn Pro Thr Ala Phe Leu Glu Arg Leu Arg	
245 250 255	
gag gca ttg agg aag cgt gcc tct ctg tca cct gac tct tct gaa ggc	816
Glu Ala Leu Arg Lys Arg Ala Ser Leu Ser Pro Asp Ser Ser Glu Gly	
260 265 270	
caa cta atc tta aag cgt aag ttt atc act cag tca gct gca gac att	864
Gln Leu Ile Leu Lys Arg Lys Phe Ile Thr Gln Ser Ala Ala Asp Ile	
275 280 285	
aga aaa aaa ctt caa aag tct gcc gta ggc ccg gag caa aac tta gaa	912
Arg Lys Lys Leu Gln Lys Ser Ala Val Gly Pro Glu Gln Asn Leu Glu	
290 295 300	
acc cta ttg aac ttg gca acc tcg gtt ttt tat aat aga gat cag gag	960
Thr Leu Leu Asn Leu Ala Thr Ser Val Phe Tyr Asn Arg Asp Gln Glu	
305 310 315 320	
gag cag gcg gaa cag gac aaa cgg gat taa aaa aaa ggc cac cgc ttt	1008
Glu Gln Ala Glu Gln Asp Lys Arg Asp Lys Lys Lys Gly His Arg Phe	
325 330 335	
agt cat gac cct cag gca agt gga ctt tgg agg ctc tgg aaa agg gaa	1056
Ser His Asp Pro Gln Ala Ser Gly Leu Trp Arg Leu Trp Lys Arg Glu	
340 345 350	

23

aag ctg ggc aaa ttg aat gcc taa  
 Lys Leu Gly Lys Leu Asn Ala  
           355                  360

1080

<210> 29  
 <211> 20  
 <212> PRT  
 <213> Homo sapiens

<400> 29  
 Thr Ser Phe Val Glu Lys Ala Asn Gly Val Lys Cys His Lys Tyr Lys  
   1                  5                  10                  15  
 Leu Ser Phe His  
                   20

<210> 30  
 <211> 63  
 <212> PRT  
 <213> Homo sapiens

<400> 30  
 Glu Thr Thr His Asn Tyr Val Lys Ser Val Ile Tyr Ala Leu Gln Glu  
   1                  5                  10                  15  
 Ala Phe Arg Val Tyr Leu Pro Ile Pro Ala Ser Pro Thr Pro Ser Pro  
                   20                  25                  30  
 Thr Asn Lys Asp Pro Pro Ser Thr Gln Met Val Gln Lys Glu Ile Asp  
           35                  40                  45  
 Lys Arg Val Asn Ser Glu Pro Lys Ser Ala Asn Ile Pro Gln Leu  
   50                  55                  60

<210> 31  
 <211> 79  
 <212> PRT  
 <213> Homo sapiens

<400> 31  
 Pro Leu Gln Ala Val Gly Gly Arg Glu Phe Gly Pro Ala Arg Val His  
   1                  5                  10                  15  
 Val Pro Phe Ser Leu Pro Asp Leu Lys Gln Ile Lys Thr Asp Leu Gly  
           20                  25                  30  
 Lys Phe Ser Asp Asn Pro Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu  
           35                  40                  45  
 Gly Gln Phe Phe Asp Leu Thr Trp Arg Asp Ile Met Ser Leu Leu Asn  
   50                  55                  60  
 Gln Thr Leu Thr Pro Asn Glu Arg Ser Ala Thr Ile Thr Ala Ala  
   65                  70                  75



24

<210> 32  
 <211> 21  
 <212> PRT  
 <213> Homo sapiens

<400> 32  
 Glu Phe Gly Asp Leu Trp Tyr Leu Ser Gln Val Asn Asp Arg Met Thr  
     1                    5                    10                    15  
 Thr Glu Glu Arg Glu  
                     20

<210> 33  
 <211> 142  
 <212> PRT  
 <213> Homo sapiens

<400> 33  
 Phe Pro Thr Gly Gln Gln Ala Val Pro Ser Leu Asp Pro His Trp Asp  
     1                    5                    10                    15  
 Thr Glu Ser Glu His Gly Asp Trp Cys Cys Arg His Leu Leu Thr Cys  
                     20                    25                    30  
 Val Leu Glu Gly Leu Arg Lys Thr Arg Lys Lys Ser Met Asn Tyr Ser  
                     35                    40                    45  
 Met Met Ser Thr Ile Thr Gln Gly Arg Glu Glu Asn Pro Thr Ala Phe  
     50                    55                    60  
 Leu Glu Arg Leu Arg Glu Ala Leu Arg Lys Arg Ala Ser Leu Ser Pro  
     65                    70                    75                    80  
 Asp Ser Ser Glu Gly Gln Leu Ile Leu Lys Arg Lys Phe Ile Thr Gln  
                     85                    90                    95  
 Ser Ala Ala Asp Ile Arg Lys Lys Leu Gln Lys Ser Ala Val Gly Pro  
                     100                    105                    110  
 Glu Gln Asn Leu Glu Thr Leu Leu Asn Leu Ala Thr Ser Val Phe Tyr  
                     115                    120                    125  
 Asn Arg Asp Gln Glu Glu Gln Ala Glu Gln Asp Lys Arg Asp  
     130                    135                    140

<210> 34  
 <211> 29  
 <212> PRT  
 <213> Homo sapiens

<400> 34  
 Lys Lys Gly His Arg Phe Ser His Asp Pro Gln Ala Ser Gly Leu Trp  
     1                    5                    10                    15

25

Arg Leu Trp Lys Arg Glu Lys Leu Gly Lys Leu Asn Ala  
                   20                                  25

<210> 35  
 <211> 685  
 <212> PRT  
 <213> Homo sapiens

<400> 35  
 Pro Lys Thr Ala Asn Leu Val Ala Asp Ile Thr Ser Leu Ala Lys Tyr  
   1                                  5                                  10                                  15  
 Gln Gln Val Leu Lys Thr Leu Gln Gly Thr Tyr Pro Xaa Glu Glu Gly  
                                   20                                  25                                  30  
 Lys Glu Leu Phe His Pro Cys Asp Met Val Leu Val Lys Ser Leu Pro  
                                   35                                  40                                  45  
 Ser Asn Ser Pro Ser Leu Asp Thr Ser Trp Glu Gly Pro Tyr Pro Val  
                                   50                                  55                                  60  
 Ile Leu Ser Thr Pro Thr Ala Val Lys Val Ala Gly Val Glu Ser Trp  
   65                                  70                                  75                                  80  
 Ile His His Thr Xaa Val Lys Ser Trp Ile Leu Pro Lys Glu Pro Glu  
                                   85                                  90                                  95  
 Asn Pro Gly Asp Asn Ala Ser Tyr Ser Cys Glu Pro Leu Glu Asp Leu  
                                   100                                  105                                  110  
 Arg Leu Leu Phe Lys Gln Gln Pro Gly Gly Lys Xaa Leu Lys Ser Xaa  
                                   115                                  120                                  125  
 Ile Pro Met Ala Leu Pro Tyr His Ile Phe Leu Phe Thr Val Leu Leu  
   130                                  135                                  140  
 Pro Ser Phe Thr Leu Thr Ala Pro Pro Pro Cys Arg Cys Met Thr Ser  
  145                                  150                                  155                                  160  
 Ser Ser Pro Tyr Gln Glu Phe Leu Trp Arg Met Gln Arg Pro Gly Asn  
                                   165                                  170                                  175  
 Ile Asp Ala Pro Ser Tyr Arg Ser Leu Ser Lys Gly Thr Pro Thr Phe  
                                   180                                  185                                  190  
 Thr Ala His Thr His Met Pro Arg Asn Cys Tyr His Ser Ala Thr Leu  
                                   195                                  200                                  205  
 Cys Met His Ala Asn Thr His Tyr Trp Thr Gly Lys Met Ile Asn Pro  
  210                                  215                                  220  
 Ser Cys Pro Gly Gly Leu Gly Val Thr Val Cys Trp Thr Tyr Phe Thr  
  225                                  230                                  235                                  240  
 Gln Thr Gly Met Ser Asp Gly Gly Gly Val Gln Asp Gln Ala Arg Glu  
                                   245                                  250                                  255

Lys His Val Lys Glu Val Ile Ser Gln Leu Thr Arg Val His Gly Thr  
 260 265 270  
 Ser Ser Pro Tyr Lys Gly Leu Asp Leu Ser Lys Leu His Glu Thr Leu  
 275 280 285  
 Arg Thr His Thr Arg Leu Val Ser Leu Phe Asn Thr Thr Leu Thr Gly  
 290 295 300  
 Leu His Glu Val Ser Ala Gln Asn Pro Thr Asn Cys Trp Ile Cys Leu  
 305 310 315 320  
 Pro Leu Asn Phe Arg Pro Tyr Val Ser Ile Pro Val Pro Glu Gln Trp  
 325 330 335  
 Asn Asn Phe Ser Thr Glu Ile Asn Thr Thr Ser Val Leu Val Gly Pro  
 340 345 350  
 Leu Val Ser Asn Leu Glu Ile Thr His Thr Ser Asn Leu Thr Cys Val  
 355 360 365  
 Lys Phe Ser Asn Thr Thr Tyr Thr Thr Asn Ser Gln Cys Ile Arg Trp  
 370 375 380  
 Val Thr Pro Pro Thr Gln Ile Val Cys Leu Pro Ser Gly Ile Phe Phe  
 385 390 395 400  
 Val Cys Gly Thr Ser Ala Tyr Arg Cys Leu Asn Gly Ser Ser Glu Ser  
 405 410 415  
 Met Cys Phe Leu Ser Phe Leu Val Pro Pro Met Thr Ile Tyr Thr Glu  
 420 425 430  
 Gln Asp Leu Tyr Ser Tyr Val Ile Ser Lys Pro Arg Asn Lys Arg Val  
 435 440 445  
 Pro Ile Leu Pro Phe Val Ile Gly Ala Gly Val Leu Gly Ala Leu Gly  
 450 455 460  
 Thr Gly Ile Gly Gly Ile Thr Thr Ser Thr Gln Phe Tyr Tyr Lys Leu  
 465 470 475 480  
 Ser Gln Glu Leu Asn Gly Asp Met Glu Arg Val Ala Asp Ser Leu Val  
 485 490 495  
 Thr Leu Gln Asp Gln Leu Asn Ser Leu Ala Ala Val Val Leu Gln Asn  
 500 505 510  
 Arg Arg Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly Thr Cys Leu  
 515 520 525  
 Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val  
 530 535 540  
 Thr Glu Lys Val Lys Glu Ile Arg Asp Arg Ile Gln Arg Arg Ala Glu  
 545 550 555 560

27

Glu Leu Arg Asn Thr Gly Pro Trp Gly Leu Leu Ser Gln Trp Met Pro  
 565 570 575  
 Trp Ile Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile Ile Leu Leu Leu  
 580 585 590  
 Leu Phe Gly Pro Cys Ile Phe Asn Leu Leu Val Asn Phe Val Ser Ser  
 595 600 605  
 Arg Ile Glu Ala Val Lys Leu Gln Met Glu Pro Lys Met Gln Ser Lys  
 610 615 620  
 Thr Lys Ile Tyr Arg Arg Pro Leu Asp Arg Pro Ala Ser Pro Arg Ser  
 625 630 635 640  
 Asp Val Asn Asp Ile Lys Gly Thr Pro Pro Glu Glu Ile Ser Ala Ala  
 645 650 655  
 Gln Pro Leu Leu Arg Pro Asn Ser Ala Gly Ser Ser Xaa Ser Gly Arg  
 660 665 670  
 Arg Pro Thr Ser Pro Thr Ala Leu Arg Phe Ser Cys Xaa  
 675 680 685

<210> 36  
 <211> 360  
 <212> PRT  
 <213> Homo sapiens

<400> 36  
 Thr Ser Phe Val Glu Lys Ala Asn Gly Val Lys Cys His Lys Tyr Lys  
 1 5 10 15  
 Leu Ser Phe His Xaa Glu Thr Thr His Asn Tyr Val Lys Ser Val Ile  
 20 25 30  
 Tyr Ala Leu Gln Glu Ala Phe Arg Val Tyr Leu Pro Ile Pro Ala Ser  
 35 40 45  
 Pro Thr Pro Ser Pro Thr Asn Lys Asp Pro Pro Ser Thr Gln Met Val  
 50 55 60  
 Gln Lys Glu Ile Asp Lys Arg Val Asn Ser Glu Pro Lys Ser Ala Asn  
 65 70 75 80  
 Ile Pro Gln Leu Xaa Pro Leu Gln Ala Val Gly Gly Arg Glu Phe Gly  
 85 90 95  
 Pro Ala Arg Val His Val Pro Phe Ser Leu Pro Asp Leu Lys Gln Ile  
 100 105 110  
 Lys Thr Asp Leu Gly Lys Phe Ser Asp Asn Pro Asp Gly Tyr Ile Asp  
 115 120 125  
 Val Leu Gln Gly Leu Gly Gln Phe Phe Asp Leu Thr Trp Arg Asp Ile  
 130 135 140

28

Met Ser Leu Leu Asn Gln Thr Leu Thr Pro Asn Glu Arg Ser Ala Thr  
 145 150 155 160  
 Ile Thr Ala Ala Xaa Glu Phe Gly Asp Leu Trp Tyr Leu Ser Gln Val  
 165 170 175  
 Asn Asp Arg Met Thr Thr Glu Glu Arg Glu Xaa Phe Pro Thr Gly Gln  
 180 185 190  
 Gln Ala Val Pro Ser Leu Asp Pro His Trp Asp Thr Glu Ser Glu His  
 195 200 205  
 Gly Asp Trp Cys Cys Arg His Leu Leu Thr Cys Val Leu Glu Gly Leu  
 210 215 220  
 Arg Lys Thr Arg Lys Lys Ser Met Asn Tyr Ser Met Met Ser Thr Ile  
 225 230 235 240  
 Thr Gln Gly Arg Glu Glu Asn Pro Thr Ala Phe Leu Glu Arg Leu Arg  
 245 250 255  
 Glu Ala Leu Arg Lys Arg Ala Ser Leu Ser Pro Asp Ser Ser Glu Gly  
 260 265 270  
 Gln Leu Ile Leu Lys Arg Lys Phe Ile Thr Gln Ser Ala Ala Asp Ile  
 275 280 285  
 Arg Lys Lys Leu Gln Lys Ser Ala Val Gly Pro Glu Gln Asn Leu Glu  
 290 295 300  
 Thr Leu Leu Asn Leu Ala Thr Ser Val Phe Tyr Asn Arg Asp Gln Glu  
 305 310 315 320  
 Glu Gln Ala Glu Gln Asp Lys Arg Asp Xaa Lys Lys Gly His Arg Phe  
 325 330 335  
 Ser His Asp Pro Gln Ala Ser Gly Leu Trp Arg Leu Trp Lys Arg Glu  
 340 345 350  
 Lys Leu Gly Lys Leu Asn Ala Xaa  
 355 360

<210> 37  
 <211> 26  
 <212> ADN  
 <213> Homo sapiens

<400> 37  
 ggaccataga ggacactcca ggacta

26

<210> 38  
 <211> 25  
 <212> ADN  
 <213> Homo sapiens

<400> 38  
cctcagtcct gctgctggat catct 25

<210> 39  
<211> 27  
<212> ADN  
<213> Homo sapiens

<400> 39  
cctccaagca gtgggaggaa gagaatt 27

<210> 40  
<211> 28  
<212> ADN  
<213> Homo sapiens

<400> 40  
ccttcctgt gttattgtgg acatcatt 28

<210> 41  
<211> 30  
<212> ADN  
<213> Homo sapiens

<400> 41  
ggaagaagtc tatgaattat tcaatgatgt 30

<210> 42  
<211> 27  
<212> ADN  
<213> Homo sapiens

<400> 42  
gggacacaga atcagaacat ggagatt 27

<210> 43  
<211> 27  
<212> ADN  
<213> Homo sapiens

<400> 43  
gccttcagaa gagtcagggtg acagaga 27

<210> 44  
<211> 25  
<212> ADN  
<213> Homo sapiens

<400> 44  
gagcctccaa agtccacttg cctga 25

<210> 45  
<211> 29  
<212> ADN  
<213> Homo sapiens

<400> 45  
gatttcagta tctactagtc tgggtagat

29

<210> 46  
<211> 27  
<212> ADN  
<213> Homo sapiens

<400> 46  
ctaggaaatc cagctagtcc tgtctca

27

<210> 47  
<211> 28  
<212> ADN  
<213> Homo sapiens

<400> 47  
ccaagacagc caacttagtt gcagacat

28

<210> 48  
<211> 28  
<212> ADN  
<213> Homo sapiens

<400> 48  
ggacgctgca ttctccatag aaactctt

28

<210> 49  
<211> 29  
<212> ADN  
<213> Homo sapiens

<400> 49  
gcaatactac atacacaacc aactcccaa

29

<210> 50  
<211> 26  
<212> ADN  
<213> Homo sapiens

<400> 50  
gggggaggca tatccaacag ttagta

26

31

<210> 51  
<211> 30  
<212> ADN  
<213> Homo sapiens

<400> 51  
ccatctacac tgaacaagat ttatacactt 30

<210> 52  
<211> 28  
<212> ADN  
<213> Homo sapiens

<400> 52  
aatgccagta cctagtgcac ctagcact 28

<210> 53  
<211> 31  
<212> ADN  
<213> Homo sapiens

<400> 53  
cgaatacaac gtagagcaga ggagcttcga a 31

<210> 54  
<211> 28  
<212> ADN  
<213> Homo sapiens

<400> 54  
agcccaagat gcagtccaag actaagat 28

<210> 55  
<211> 27  
<212> ADN  
<213> Homo sapiens

<400> 55  
gcgtagtaga ggttgtgcag ctgagat 27

<210> 56  
<211> 27  
<212> ADN  
<213> Homo sapiens

<400> 56  
cccttaccaa gagtttctat ggagaat 27

<210> 57  
<211> 27  
<212> ADN  
<213> Homo sapiens



<400> 57  
accgctctaa ctgcttcctg ctgaatt

27

<210> 58  
<211> 420  
<212> PRT  
<213> Homo sapiens

<400> 58  
Thr Ser Phe Val Glu Lys Ala Asn Gly Val Lys Cys His Lys Tyr Lys  
1 5 10 15  
Leu Ser Phe His Xaa Glu Thr Thr His Asn Tyr Val Lys Ser Val Ile  
20 25 30  
Tyr Ala Leu Gln Glu Ala Phe Arg Val Tyr Leu Pro Ile Leu Pro Ala  
35 40 45  
Ser Pro Thr Pro Ser Pro Thr Asn Lys Asp Pro Pro Ser Thr Gln Met  
50 55 60  
Val Gln Lys Glu Ile Asp Lys Arg Val Asn Ser Glu Pro Lys Ser Ala  
65 70 75 80  
Asn Ile Pro Gln Leu Xaa Pro Leu Gln Ala Val Gly Gly Arg Glu Phe  
85 90 95  
Gly Pro Ala Arg Val His Val Pro Phe Ser Leu Pro Asp Leu Lys Gln  
100 105 110  
Ile Lys Thr Asp Leu Gly Lys Phe Ser Asp Asn Pro Asp Gly Tyr Ile  
115 120 125  
Asp Val Leu Gln Gly Leu Gly Gln Phe Phe Asp Leu Thr Trp Arg Asp  
130 135 140  
Ile Met Ser Leu Leu Asn Gln Thr Leu Thr Pro Asn Glu Arg Ser Ala  
145 150 155 160  
Thr Ile Thr Ala Ala Xaa Glu Phe Gly Asp Leu Trp Tyr Leu Ser Gln  
165 170 175  
Val Asn Asp Arg Met Thr Thr Glu Glu Arg Glu Xaa Phe Pro Thr Gly  
180 185 190  
Gln Gln Ala Val Pro Ser Leu Asp Pro His Trp Asp Thr Glu Ser Glu  
195 200 205  
His Gly Asp Trp Cys Cys Arg His Leu Leu Thr Cys Val Leu Glu Gly  
210 215 220  
Leu Arg Lys Thr Arg Lys Lys Ser Met Asn Tyr Ser Met Met Ser Thr  
225 230 235 240  
Ile Thr Gln Gly Arg Glu Glu Asn Pro Thr Ala Phe Leu Glu Arg Leu  
245 250 255

33

Arg Glu Ala Leu Arg Lys Arg Ala Ser Leu Ser Pro Asp Ser Ser Glu  
                   260                                  265                                  270  
 Gly Gln Leu Ile Leu Lys Arg Lys Phe Ile Thr Gln Ser Ala Ala Asp  
                   275                                  280                                  285  
 Ile Arg Lys Lys Leu Gln Lys Ser Ala Val Gly Pro Glu Gln Asn Leu  
                   290                                  295                                  300  
 Glu Thr Leu Leu Asn Leu Ala Thr Ser Val Phe Tyr Asn Arg Asp Gln  
                   305                                  310                                  315                                  320  
 Glu Glu Gln Ala Glu Gln Asp Lys Arg Asp Xaa Lys Lys Gly His Arg  
                                   325                                  330                                  335  
 Phe Ser His Asp Pro Gln Ala Ser Gly Leu Trp Arg Leu Trp Lys Arg  
                                   340                                  345                                  350  
 Glu Lys Leu Gly Lys Leu Asn Ala Xaa Xaa Gly Leu Leu Pro Val Arg  
                   355                                  360                                  365  
 Ser Thr Arg Thr Leu Xaa Lys Arg Leu Ser Lys Xaa Lys Xaa Ala Ala  
                   370                                  375                                  380  
 Pro Ser Ser Met Pro Leu Ile Ser Arg Glu Ser Leu Glu Gly Pro Leu  
                   385                                  390                                  395                                  400  
 Pro Gln Gly Thr Lys Val Leu Xaa Val Arg Ser His Xaa Pro Asp Ser  
                                   405                                  410                                  415  
 Ser Ser Arg Thr  
                                   420

<210> 59  
 <211> 32  
 <212> ADN  
 <213> Homo sapiens

<400> 59  
 taaactacaa atggttcttc aaatggagcc ca

32

<210> 60  
 <211> 32  
 <212> ADN  
 <213> Homo sapiens

<400> 60  
 gatgcagtcc aagatgcagt ccatgactaa ga

32

<210> 61  
 <211> 1740  
 <212> ADN  
 <213> Homo sapiens

&lt;400&gt; 61

```

aggttggtg acaaccgctc ttaactgctt catgctgaat tggggcatag taggggtcgt 60
gcagttgaga tttccttggg aggggtgcct tcaatgtcat caacattgga gcatgggcta 120
gcaggccagt ccagggtgct gcggtagatc ttagtcatgg actgcatctg gggctccatt 180
tgaagaacca tttgtagttt tacagcttcg attctggaag agacaaacgt aacaaggagg 240
ttaaagatac aaggattgaa atgtacggcc tgaagtgcag gggcatatga gtgtgggagg 300
tgcaagtggg gtttccttta gaaaaactcc gatacaatag ggcatacaata tttctaggaa 360
gccacattct ccatagaagc tctcggttaag gggagctact ggtagtacag cagcatacag 420
gggtgagcgt gagagtgaag gggggttaaga gaacagtaaa aagaaaaata tgacaaggga 480
gggccaagag gatctacgat tctagtact ttctcacgg ttgtcgcctg aagagcaggc 540
gcagatcctc tagaggttca caggaaatagc tagcattgtc tgctggattt tcgggttcct 600
ttggcagtat ccagggtttg gctcgagtgt gacttatcca agactccact ccagccactt 660
aactgcggtt agggtagata aaatgactgg gtagggctct tcccaggatg tgtgtaggga 720
tggggaattc aaggggaagg gacttgacta ataccatgtc accagggtgg aataattcct 780
ttccctctc tcagggacag gttccctgta atgttttaag aactcgttga tatttggcta 840
aggaggtgat gtctgcaact aagttggcgg tctctcagtc aagcacaagg tcattgggta 900
ggaagggtcg tccatacagc atctcatatg gactaagtcc tgcttttttg ggacagtttc 960
ggattcttag taaggctata ggcaacagag caggccatgc aagggtgggtt tcttgggta 1020
gcttttttag atgtcgtttg agtggttcat tcattttctc aacttttctt gaggatcgtg 1080
gcctccaggg acagtgttaag tgatattgta tacctaaccg ctgggatact ccctgcgtta 1140
ctgcagcctt gaaattgggg ccattgtcac tctgtaaacc tcagggaagt ccgaatctgg 1200
gaattatttc atgaattagt acttttatta cctcttgggc cttttctgtc ctacaaggga 1260
aggcctccac ccaaccagtg aaagtaccca gattagtaga tactgaaatc tctgagattt 1320
gggcatgtgg gtaaaatcta gttgctagtc ttctcctggg taatggcctg ttctttgttc 1380
tcctgaaggg gcttggcaat aaggcagggg attatttctt tggcacactt cacaggccct 1440
gactatctgc ttgacagttt tgaaaaggcc tggccaagta aataatgatt tggccatctg 1500
atgggtgctg tcaatgccta agtgaaaggc ctggtgaagg gttttaagta atttccattg 1560
gttagctgca ggcaaaagta ttttttcttt ggtggctggc catcctgagg agaggaaact 1620
atgtcctcgt gagtttcccc attccatttc ttctgctgag tactggagct tggtttccca 1680
gaggggatta ccccatacta ggggtccttc tgtaagcatt tctaattggag agtcctgcct 1740

```

&lt;210&gt; 62

&lt;211&gt; 7140

&lt;212&gt; ADN

&lt;213&gt; Homo sapiens

&lt;400&gt; 62

```

ttggtcttaa gaacacaaat gatatggctc caatgactgg aggaacacca gggtccttgg 60
tctcacgctg atttagataa aacgactgtc aggcctctga gcccaagcta agccatcctc 120
ccctgtgacc tgcacgtata catccagatg gcctgaagta accaaagaat cacaaaagca 180
gtgaaaatgg cctgttctct ccttaactga tgacattcca ccatttgtat ttgttctctg 240
cccatcttaa ctgagcgatt aaccttgtga aattccttct cctggctcaa aacctcccc 300
actgagcacc ttgtgacccc cgccccgcc cctaagagaa aacccccctt gattataatt 360
ttccactacc caccacaaatc ctataaaatg gccccacccc tatctccctt cgctgactcc 420
tttttcggac tcagcccggc tgcacccagg tgaaataaac agccttgttg ctcacacaaa 480
gcctgtttgg tggactctct tcacacggac gctcatgaca tttggtgcca aaacctggga 540
taggaggact ccttcaggag accagtcccc tgtccttgcc ctactctgt gaggacatcc 600
acctaacaac ttgggtcctc agaccaacca gcccaaggaa cagctacca atttcaaatc 660
aggtaagcag tcttttctct ctcttctcca gcctctcttg ctacccttca aactccctct 720
ctcactacc ttcaatctcc ctgtccttcc aattccagtt cttttctatc tctagtagag 780
acaaaggaga cacattttat ccatggaccc aaaactccag caccagtcac ggacttggga 840
agacagtctt cccttgggtg ttaatcactg cggggacgcc tgctgatta ttacccaca 900
ctccattggt gtctgatcac ggtggggaca cctgccttgg tcaactaccc acattccctt 960
ggtggtacgt caactgcaa agcaggggac gcctgcttgg gctgtcacc caccctctt 1020
tctgtgtctc tacctttctc tttaaactta cctccttcac tatgggcaa cttctgcctt 1080
ccattcccc ttcttctccc tttagcctgtg ttcttaaaaa cctaaaacct cttcaactca 1140

```

cacctgacct	aaaaccta	aa	tgcttatt	tcttctgca	cactgctgg	ctgcagtaca	1200
aacttgataa	tagctttaa	tg	gccagaat	atggcacttt	caatttctcc	atcctacaag	1260
atctagataa	tttttggtg	aaa	atgggaaa	aatggtctga	gatgctgac	gtccaggca	1320
tcttttacac	attggtccct	cc	tagtctc	tgctcccaat	gcgactcatc	ccaaatcttt	1380
cttctttctc	tcctgtctgt	tc	cttcagtc	tccaccccaa	gctctgagtc	ctttgaatcc	1440
tcctttgcta	cagacccatc	tg	aactctcc	cctcctcccc	aggctgctcc	tcaccaggcc	1500
gagccaggtc	ccaattcttc	ct	cagcctct	gctccccac	cctataatcc	ttttatcacc	1560
tcctctcctc	acactcagtc	cg	cttacag	tttcgttctg	tgactagccc	tccccatct	1620
gccccacaat	ttcctcttaa	ag	aggtggct	ggagctaaag	gcatagtcaa	ggttaatgct	1680
cctttttctt	tatctgacct	ct	cccaaatc	agttagcggt	tacgctcttt	ttcatcaaat	1740
ataaaaaccc	agccagttca	tg	gccatct	ggcaacaacc	cttacaggct	ttacagccct	1800
agaccctgaa	gggtcagaag	gcc	gtcttat	tctcaatatg	cattttatta	cccaatccgc	1860
tcccaacatt	aaataaagct	cc	aaaaatta	aattctggcc	ctcaaacccc	acaacaggac	1920
ttaattaacc	tcacttcaag	gt	gtacaaga	atagagtaga	ggcagccaag	tagcaacgta	1980
tttgagttgc	aattccttgc	ct	caactctg	agagaaaccc	cagccacatc	tccagcaaac	2040
aagaacttca	aaacacctga	act	gcagcag	ccaggcggtc	ctccaggacc	acctcccca	2100
ggatcttgct	tcaagtgcgc	gaa	atctgac	cattgggcca	aggaatgcct	gcagcccagg	2160
attcctccta	agccacgtcc	cat	ttgtgca	ggacccact	ggaaatcgga	ctgtccaaact	2220
caccgggacg	ccaatccag	ag	cccttga	actctggccc	aaggctctct	gactgactcc	2280
ttcccgatc	ttctcggtct	ag	cagctgaa	gactgacact	gcccgatcac	ttcagaagtc	2340
ccctggacca	tcacggatac	tg	agcttcag	gtaactctca	cagtggaggc	taagtccatc	2400
ccctgtttaa	tcgatacagg	gg	ctaccac	tccacatcac	cttcttttca	agggcctggt	2460
tccttttccc	ccataactgt	tgt	gggtatt	gacggccaag	cttcaaaacc	ccttaaaact	2520
ccccactct	ggtgccaact	tg	gacaacat	tcttttatgc	actcttttcc	agttatcttc	2580
acctgcccag	ttcccttatt	agg	ccgagac	attttaacca	aattatctgc	ttccccgact	2640
attcctgggc	tacagccaca	tct	ccttgcc	gcccttcttc	ccaacccaaa	gcctccttca	2700
tatcttctct	tcatatcccc	cc	accttaac	ccacaagtat	gggacacctc	tactccctcc	2760
ctggcaaccg	atcacacgcc	cat	tactatc	ccattaaaac	ctaatacccc	ttaccttgct	2820
caatgccagt	atcccatacc	aca	acaggct	ttaaagggat	tgaagcctgt	tatcacttgc	2880
ctgtctacag	acgggcttct	aaa	acctata	aactctccat	acaattcccc	cattttacct	2940
gtctaaaaac	cagataagtc	tt	acaggtta	gttcagaatc	tgcaccttat	caaccaaat	3000
gttttgccca	tcaccctgt	ag	cacccaac	tcgtacactc	ttttgtcctc	aatgccttcc	3060
cccacaactc	actattccgt	tct	tgtatctt	aaagatgctt	ttttcactat	tcccctgcac	3120
ccctcatccc	agcctctctt	tg	cttttacc	tggtactgac	ctgacaccca	tcagtcccag	3180
cagcttacct	gggctgtact	gcc	gcaaggc	ttcagggaca	gccctcatta	cttcagccaa	3240
gctcttctct	atgatttact	tt	ctttccac	ctctctgctt	ctcaccttat	tcaatatatt	3300
gatgaccttc	tactttgtag	ccc	ctcttt	aaatcttctc	aacaagacac	cctcctgctc	3360
cttcaacatt	tgttctccaa	agg	atatcgg	gtatccccct	ccaaagctca	aatttcttct	3420
ccatctgtta	catacctcgg	cat	aattctt	catgaaaaaca	catgtgctct	ccctgccaat	3480
tgcttctcca	actgatctct	caa	atcccaa	cctcttctac	aaaacaacaa	ctcctttccc	3540
tcctaggcat	ggttggtatc	ttt	tgcttt	ggatacctgg	ttttgccatc	ctaacaaaat	3600
cattatataa	actcacaana	gg	aaacctag	ctgaccccat	agattctaaa	tcctttcccc	3660
actcctcttt	ccattccttg	aag	acagctt	tagagactgc	tcccacacta	gctctccctg	3720
tctcatccca	acccttttca	tt	acacacag	ccgaagtgca	gggctgtgca	gtcggaaatc	3780
ttacacaagg	accgggacca	tg	ccctgtag	cctttttgtc	caaacaactt	gaccttactg	3840
ttttaggctc	gccatcatgt	ct	ccatgcgg	tagcttccgc	tgccctaata	cttttagagg	3900
ccctcaaaat	cacaaactat	gct	caactca	ctctctacag	ctctcacaac	ttccaaaatc	3960
tattttcttt	ctcacacctg	ac	gcatatac	tttctgtctc	ccggctcctt	cagctgtatt	4020
cactctttgt	tgagtctccc	aca	attacca	ttcttctctg	cccagacttc	aatctggcct	4080
cccacattat	tctggatacc	aca	acctgacc	ctgatgattg	tatgtctctg	atctacctga	4140
cattcacccc	atttcccat	att	tccttct	tttctgttcc	tcagtgtgat	cacatttggt	4200
ttactgagcg	cagttccacc	agg	ctgac	gccactcacc	agcaaaaggca	ggctatgcta	4260
tagaatcttc	cacatccatc	att	gaggcta	ctgctctgcc	cccctccact	acctctcagc	4320
aaagccgaact	gattgcctta	act	cgggcct	tactcttgc	aaagggacta	cacgtcaata	4380
tttatactga	ctctaaatat	gc	cttccata	tcttgaccaa	ccatgctgtt	atatgggctg	4440
aaagagggtt	cctcactacg	caa	gggtcct	ccatcattaa	tgctcttcta	ataaaaaactc	4500
ttctcaaggc	tgctttactt	cca	aagggaag	ctggagtcac	acactgcaag	ggccacacaa	4560
aggcgtcaga	tcccattact	ct	aggaaatg	cttatgctga	taaggtagct	aaagaagcac	4620

```

ctagcggtcc aacttctgtc cctcatggcc agtttttctc cttcccatca gtcattccca 4680
cctactcccc cattgaaact tccgcctatc aatctcttct cacacaaggc aaatggttct 4740
tagaccaagg aaaatatctc cttccagcct cacaggccca ttctattctg tcatcatttc 4800
ataacctctt ccatgtaggt tacaagccac tagtccacct cttagaacct ctcatttctt 4860
tccatcggtg aaacatatcc tcaaggaaat cacttctcag tgttccatct gctattctac 4920
taccctcag ggattgttca gggccctccc cctccctaca catcaagctc ggggatttgc 4980
ccctgcccag gactggcaaa ttgactttac tcacatgccc tgagtcagga aactaaaaata 5040
cctcttggtc tgggtagaca ctgtcactgg atgggtagag gcctttccca cagggtctga 5100
gaaggccact gcagtcattt cttcccttct gtacagacata attccttggg ttggccttcc 5160
cacctctata cagtccaata acggagcagc ctttattagt caaatcacct gagcagtttt 5220
tcaggctctt ggtattcagt ggaaccttcg taccctttac tgcctcctcaat cttcaggaaa 5280
ggtagaatgg actaatggtc ttttaaaaac acacccacc aaactcagcc tccaacttaa 5340
aaaggaggat agagcccaaa aactcgcaac caagctagta attatgctga accccttgg 5400
gcactctcta attggatgtc ttaggtcttc ccaaatctta gtcctttaat atctgtttt 5460
cctcttctct tactcgacc ttgtgtcttc cgtttagttt ttcaattcat acaaaaacgc 5520
atccaggcca tcaccaatcg ttctatacaa taaatgctcc ttctaacaac cccacaatat 5580
cgccctttac cacaaaaatct tccttcagct taatctctcc cactctaggt tcccatgccg 5640
cccataatcc ctctcgaaag agccctgaga aacatagccc attatctctc cataccacc 5700
ccaaaatttt ggtgcccaca acacttcaac actattttac attatttttc ttattaatat 5760
aagaagacag caatgtcagg cctctgagcc caagccatca tatccctgt gacctgcaca 5820
tatacatcca gatggcctga agtaactgaa gaatcacaaa agaagtgaag atggcctgtt 5880
cctgccttaa ccgatgacat tccaccactg tgatttgttc ctgccccacc ttaactgagc 5940
aattaacctt gggaaattcc ttctcctggc tcaaaaacct cccactgag caccttgta 6000
ccctgcccc tccactacc acccaaatcc tataaatgg cccacccca tctcccttag 6060
ctgactcctt ttttggtc agcccgctg caccaggtg aaataaacag ccttgttgc 6120
cacacaagc ctgtttggtg gactctcttc acagggacgg gggtgacaac aacacggaca 6180
cacattggat ggttttaagg agcagagagt ttaatacgca aaaaagaagg aagaggctcc 6240
cctgtacaga cacagagga gggggtccca agccgagaga aggaaccccc atgtgcagt 6300
gaaaagtggg tgattatact gggaggctgg aggaggcggt gtctgatttg cacaggggcc 6360
aggggattgg gttgaccagg tgtatcatc atgtaccccg caaaaaacct ggccctccca 6420
cctcagccct ttaatatgca aatgtgggtt gccatgatgt tctgaaaaca catgaattat 6480
ctggaggggg ccatgacact tggtagatgt gctgacaaga agaggggtgg aatcgccatg 6540
gtggccatgt tgggtggacc tagtttttaa tagcctgcat ttgcatatca aagtttgctg 6600
gcctggctct ttaagctgtc ttttctgtta gaaaaggaat ggtttggaat ggggtgaggt 6660
tgcttcttat tacaagaaaa ttccaaaaaa ctttactct ttctagctgc caaaaaacta 6720
tttcttaata acttatgtat taccataatt aggcagcacc aaagatccct gcaggtcaga 6780
ccactgcaat taacatgctg gctttactgc tgattatgg agctgcatcc acctagcctc 6840
tcattatgca actgcctgac ctctgccacc ccacgagcca cttatcccca cttataatca 6900
gccatttctg attgtaacat ctgccactta ttcccgacgt tgtggtatat cctatagatg 6960
aattcattca acatccattc caacaccacc tctcttgctt tcctatactc tctggagagt 7020
gaattactga gtcacatgat cttcactgca gtcatttctg gctatgtgac atagtctctg 7080
acagtgaaca tagacagaag tccctggggc gggcttctct tctgggatga gggcaaacg 7140

```

&lt;210&gt; 63

&lt;211&gt; 44100

&lt;212&gt; ADN

&lt;213&gt; Homo sapiens

&lt;400&gt; 63

```

tgcctttatt tccgtaggct ggtcatatgg cgctagcact cacataaagc taccgaggag 60
agcgaatgaa accaaaaatca ctttaccttc acagcacgag gccgtcgtcc ctctcgatat 120
ttggcccggtg tgcgcataac cgccctctgg acgtggtgat caaataaaact ccctagctcc 180
ccgcccgtctg acgccatctt gcctactttg atcctcgag ggaggacaac atccgcctta 240
ctgagctccc ttttatccaa taagagagcg ggatgagtta aggagtgcga ggattggctg 300
gagaatcgac agcgtcggcc atcgtttcct gcgtgcgaag atttgatgaa cgagggtgcc 360
ccccgagcg gctcggcgga gaggcgcggt ggggtgacaga agctttcttg tcccaccac 420
tacaggctta cggcaggatg cgacgagggg agaggggcg gggccgagg gggcgggg 480

```

gacgatctc	ctccggctcc	gacgtcctcg	gcctgccggg	tcccgggtcc	tttgcggcgc	540
taggggtggc	gaacccagag	cgacgctccg	ggacgatgtg	gggcagcgat	cgccctggcg	600
gtgctggggg	aggcggggcg	gcagtgactg	tggccttcac	caacgctcgc	gactgtctcc	660
tccacctgcc	gcggcgctctc	gtggcccagc	tgcacttgct	gcaggtaacc	tgccggcccc	720
gagccacctg	atcttcagcc	tggggtcgga	cgaggccgaa	gcctctcagg	gacgcggcgg	780
gacaccggct	gccacccggg	cgccgcccga	gcgcgcagag	atcagggtcc	ctcgacggca	840
gggcccctct	gggtagtctc	tggatcccac	aagtccagtg	cagccctggg	ctcgtcttat	900
cccaggtctt	ttcacttggt	gaaactgaac	ctagaaacgt	cctaataattc	taccactgtt	960
tttataaata	ttccttattc	caggctggaa	aagctcctga	gaagtgggtt	gtttttatta	1020
ttttaaaagg	tgttttcctt	gccagccatt	tccagttaac	ctgcgctgct	gccgtccggg	1080
ccgcgagagc	gggacgcaga	gttgttggcg	gagccctgt	cggttcccgg	ggactaagca	1140
cccgctccca	tgagcgggaa	aggttaatac	aatgatggtt	ctgccctgcg	tcgctgacgc	1200
ggaacacagc	tgtagtgtgt	taggaacaca	taacgtagtt	aagatcactt	gaagctctgc	1260
gatcagtgcg	ccttctggac	gttgtgggtta	ggatgtttca	cagttctaac	cactgggtga	1320
gatacagcgt	ccatattttc	ataattaaaa	atagaggcac	atggtctcac	gagtttgagt	1380
gtactttatg	tgagcaaaag	acggcgatatt	tgaaatcctc	ataaatcctg	gatgcatggt	1440
accaccagct	ggctaatact	tgcaatgaat	agagtttgca	ataatttcaa	gcatcccttc	1500
tttccacttg	agttaacttc	ccatacctag	gggaagatat	ttttggtcca	ctgaaaacat	1560
gagttcagca	gaatcctcct	atcatcgctg	ttattatatt	ttaccactaa	gtagacaatc	1620
ttttggtttt	tgatgggctt	tatggctaga	gacaaatcag	tcactgtcac	caagttccag	1680
gtagaagttg	gttcagtgtc	ctgtcagctt	cgatgggatt	tttcaacatg	ttttcaaatc	1740
tgcacttaat	agtaggaatg	ctttcttaca	gtaactctaa	tttgatccta	agatgtagtt	1800
gttaccttac	attcatcact	gtttaagaat	ttagtggctt	tgatctttgt	tttaaatatt	1860
gagccttcgg	gaagtactta	taagaattaa	ttcatgcata	tctttttgaa	atgtaaatgt	1920
cttttagccct	ggaacaaatt	gctgtttctg	ttcagcccat	attagcagaa	taggtcaact	1980
ttactttcta	attatcaatg	taataagttt	attactttat	agattccata	aatctataca	2040
tttattcctc	gatgaattat	ataaatttat	agaattttat	ttttatagaa	aatttggaaa	2100
gcattggaaa	ttattaacaa	gaaaataaag	taccataaat	cccagaactt	agaggtagct	2160
aatgttgaca	gtttggatca	aatcttccag	ttttgtttct	aatctttatt	tttaacataa	2220
atgaggtcct	gtatacacac	gtacagtttt	gtgtcctggg	gtttttattt	aatgttatta	2280
tgagtgtttt	attttgttaa	aaggtcatca	ttttaagttg	tttaattagta	ttctagcaca	2340
aatttgccat	aatttattta	attgtttact	atgattgacc	atttagattg	tacttaattt	2400
ttaggcatta	gaagtgataa	actatatttt	aatcagacgt	tgaaaataac	acatctttgt	2460
ttagaaaaca	tcattttatt	tctggttgct	taggatagat	tcccagaatt	cttgggttag	2520
aggccataga	taattatgaa	agcagaaaag	ttcacaagtt	gggagttaat	acttgaatta	2580
cgttgtaaaa	ggtgaagcat	tgagtgcata	atacagatca	tgcatgtaag	ggaagaaggg	2640
ttggaacaat	ggttttctgg	cctatgtcag	acttaccttg	aagcttttaa	gaatacagat	2700
gttctgatca	accctcagac	ctattaaatc	agacctaaaa	tcttagggaa	taggcttttag	2760
gcatctctaa	ttttaaaaaa	tttaticagg	ctacttggat	gcacaaaaga	gttgagacct	2820
actgtcctag	aatcatagaa	ttttaatgac	gatagagacc	ttaagcatct	aggtcgtttc	2880
tgtactttta	catgtaagga	aactggcatt	cctaggccag	taccattgcc	atgcagctaa	2940
tttgccctct	tgtctatagc	tcactctgca	tcaccaaac	taccgttctc	actgtttctt	3000
ctataaccaa	tctccttccc	acttctgttc	tcttactcat	gccattcttc	cctcagtcac	3060
ttttcttctt	tccatacaaa	ttccatgtct	ttaaaaagga	ataatcctac	ctcctccaca	3120
tagctttcca	attctctgtt	gcccacattt	gtctcccttt	caatacttct	ctgttgtgtt	3180
atgtgacaca	tcacatttga	tatactctgt	actgtgtttc	aagtattgta	ttctcttggt	3240
tactcaagtc	attattttcag	gactgactac	ccagttagatg	cttttaagtca	ggattttctca	3300
accttggcac	tgttgacatt	ttgagctgga	taattttttg	ttttgggggc	tctcctgtac	3360
atttttaagat	gttttaacagc	acccttggcc	tctatccagt	agacgcctgt	actgcctccc	3420
cctatctgtg	acaaccaaaa	aggtcttcag	acattgtcag	atgtctactg	aaggacaaaa	3480
tcacctctgg	ttgagaacca	ccgcttcaac	taagttatct	tctctgtact	cagaacttga	3540
tgtgattgca	gcagggggag	aggattcata	tacacagtga	atgcaaacga	acctaaatca	3600
ccattcggat	atggccacac	aattttcatt	tcccttgtgt	tagcaagaga	taccctaggc	3660
tttggaacctg	attattccta	aggcattctg	atgtatgggt	ttacctgcag	atttcctggt	3720
aatactgata	cctcagtttg	gggtcaagaa	gggtcaattaa	ttgattgatt	tgatttgact	3780
cctggaaaag	acgtcccttt	ctagctgtct	ctttcttctc	ttacctgaa	tagccagggc	3840
tctgtggttc	aagtgaagta	ttttgacata	aaaattaaact	tagaacattg	gtctgcagag	3900
tttgctcaat	ataactgagc	acatattgtg	gctttatgga	gctggttact	actttttgac	3960

```

caaataaata attagaagta tttttcctcc tcaataaggt tcatttttcc ttttttcagt 4020
gagctggtag agtttccttt tttgatattt cagggcatct ttcataattc catctcttaa 4090
gtttcttcat atgaagtaga atttatctgg attatgtaatt gctgactctg atgaaaaccc 4140
atagaaaagca tctggggcctt gatcaccttc attcttgtaa tagctcacac ggttacagc 4200
gatatggtaa cttaagactt ttgattccaa atctaggcaa aatacactca gttgaaagaa 4260
tttgtcagcc agaacagttg gactgttctg tgaaaattgt gagaaaaatt acacaactaa 4320
gtgatacatg atgatggctt tcttaaatat aaaattgtaa taacatgggt aatttccagt 4380
acgttatatt gtcccagaag tggctccaac attgtttgaa atttgtctca tttaaagaaa 4440
cataagctgg ctatggtggc tcacgcctgt aatcccagca ctgtgggagg ctgaggcagg 4500
cagatcacct gaggtcagga gttcagagcc agcctggcca acatgggtaa accccatctc 4560
tactaaaaat acaaaaatta gccgggcatt tgggtggggc ctgtaatccc agctacttgg 4620
gaggtcaggg caggagaatt gcttgaatct gggagggtga ggttgcagt agccgagat 4680
gtgccactgt cctccagcct ggggtgacaga gtgagctctc gtctcaagaa aaaaaaaaaa 4740
aaaagcaaga aacataaaga ctgggcatgt tggctcatgc ctgtaatccc agcactttga 4800
gagactgagg tgggaagatc acttgagccc aggagggtta ggctgcagt agccgtgat 4860
ttgccactgt actcgagcct gggcaacaca gtgagatcct gtctcaggaa aaaaaaaat 4920
gcactgtaaat gaatgaattt gatatttaatt attttaaatt atgaaaactg ttctgtagag 4980
atgtagatct tgccatgttg cccaggctgg ctttgaactt ctgggctcaa acaatcctcc 5040
tgtctcagtc tcccaaagta taaagattac acatgtgagc cactgcacct ggcctaatat 5100
ttttaactta atgaatttat tttgatataa ataaattaat aacactgaag ctctctgata 5160
taataagtct tttgtgtgt gtgacgggtt ctactctgt tggccagact ggagtgtaat 5220
ggcactatca tggctcactg tagcctcaac ctccctgact caagtgatcc tcccacctcg 5280
gcttctctgag tagatgggac cacaggcgta tgccaccaca cctggctgat ttttaaaat 5340
tattattgat acatattaat aaaattattt ttattttaaa aatgatatat gtggctgggc 5400
atggtggctc atgcctgtaa tcccgcagct ttgggaggcc gaggtgggag gatcacttga 5460
gaccaggagc ttaagaccag cctaagcaac atagttagat cccatctcta tagaaaaaaa 5520
aaatggctag gtgtggtggt gtatgcctat attcccagct actcaggaga ctgaggtgag 5580
aggattgcta gagcccagga gtttcaagtt acagtgcct atgattgtgc cagtgcactc 5640
cagcctgggc aacagagcaa aatcctgtct caaaaaaaaaa aaaagttcga aaatgcttat 5700
gatgcaatat aagttagtga aaaggatatt aaattgtgcc tatatgaaca caactatatg 5760
aaaaacttgc acatagagaa aaggattaac aagaaataga ccaaattgtt cacatggttg 5820
tcttgtttgt ggagagaata tcagtagttc atttgtttcc ttccaagttt atatgttttc 5880
cgaggctctt taatgagtt tgtaattgtt taatcataga aaacctttt ttggtccttg 5940
gccacaaact tacatgtttt aatgtaattg cttttttaat gagaataaat gttatatttt 6000
gcttttttaa aacctatatt cccatagtta tatgagccct tacaattatt aagaggtgc 6060
ataataatac gtttctggaa gggtagacaa gaaacagcag taattacctc tgagaacaga 6120
gacatggctt cacattttac ccttttgtag gttttgtgt tttgccacat gcattttata 6180
ttcttccaat aaataagtaa ataaatatgg attgtatact ccactgtggt ggtgtttcat 6240
aattctaaaa ttatattgct acatttttaa agatgatatg tgtttctact tattaacgta 6300
tatgttaaaa tagtaaatat atattctatt taataatttc cctattgata gacatttaag 6360
acagctcaca gtgttcacta tcatagaaaa tactgcacag atagcttttg ctatagtttc 6420
ttttttcttt gaatcgtaa ttgggaataa atgctcaaat agttatatgt ggctcaactg 6480
ctattttaagt ttattgactg actgtgccta ttttgaattc tgaaggggtt gattaaattt 6540
ataatgctgc cataagaata taagggtatt ggcttcatta gcatccacca gcattgggtg 6600
ttggaaaatga ttatagattt ttaaattgcta caacaaatgt agataacaga gaactatcta 6660
tagaactctt tttggacatg tgaattgtaa taatagttaa ttttcatgtg aatccagaaa 6720
aatgtatacg aaaacctttt ttctctctcat ttcttatatg aatagaatca agctatagaa 6780
gtggtctgga gtcaccagcc tgcattcttg agctgggtgg aaggcaggca ttttagtgat 6840
gggggacagg taagcacatg tgatggcaat aactttcttc taatatcaca taatatagca 6900
atagaaataa aattaaaagt tttagatttt tgtaaaggga ggtgagatgt cacctaattt 6960
gtatgctatt atgtaactag tctaggatat tgaagctgac tatactctgt ttttaggtca 7020
ttatcttgta gtttaccata ctccctactt gcttcttatt ctactattta actcattttc 7080
cacatccctc aattttggtt tcatgaaatt atttttctt ctgaattact aggttctact 7140
tactattatt aaactttatt tctgacatat tttataacct tccatgggtc cacttgatta 7200
aaaataaaaa attcagctgg gtgcgggtgg tcacacctat aatcccagca ctttgggagg 7260
ccaagggtgg cggataattt gaggtcagga gttggagacc agcctgccca acgtggtgaa 7320
accccccctc tctactaaaa attcaaaaat tagctgggca tgggtggcagg tgcctgtaat 7380
cccagctact caggaggctg aggcaggaga attgcttgaa cctgggagggt ggaggttgca 7440

```

```

gtgagctgag attgcactgc tgcacttcag ctgggtgaca agagcgaaac aatgtcttga 7500
aaaaaaataa aaaataaaaa attctacaac acaggggttat tatttttcca tttttgtttt 7560
cccttatgag tttaatatgt ttagattata aacctgaaag cttgaatacc tatgtctatc 7620
ttttgttttc ttatgtttat caagtatttc cttaaacad tttctaaact gtaagaataa 7680
tgtgaggctg ggctcaatgg cttatgcctg taatcccagt gctttgggag gccaaaggtg 7740
gaggaccact tgaggccacg agttcaagat tagcctggct agccaacata gcaagaccct 7800
atctctataa aaaaattaaa aaaattagct gggcatggta gcaaatgctt gtagtcccag 7860
ctactcagca gactgaggta ggaggaatgc ttgagaccag gaatttgagt gacctatgat 7920
tatgcactcc agcccgggca atagcaagac cctatctctt aaaagaagaa gatgtagtaa 7980
taatacatat tcattataac tattttacca ttgaaagtaa aaaatgagtt tttacctttt 8040
cccagtcaca tcttcagaat ggggatctca gtagaccttt aggattggaa gaatgagatc 8100
attcatatatt ttgcaatta ttacccaca aaatatttca gatacctttc catgtattac 8160
aaacaatgtg catttaacat gtctctctct ttctctctct ctctgtgtgc gtcttcatga 8220
tcctctgttg cagccctgcc agtaagacac tatctcctga agaactcactg ataggaacag 8280
aaagtggact ggctaggcca ggagtcctta gcttcttagg gggcaggagc tgctttgtgc 8340
tttctcagaa tcagatatat atgtggactg aaacatttaa aaacagaata gccaaagggt 8400
ctatacgttt aaaacttata tagatggggc tacattgctc tctattacta atttcccatg 8460
acaatacacg agagtgccat gtcttttttaa ctgtttttga gcacagacta atcttgttta 8520
tgcatgtttt ttgatgagaa taggctactc atgagaaatc tgtaaaccta acactagtcc 8580
cttgcatact cttaattggt gctagaatct taaatattta gcaccagacg gaccttagaa 8640
atcattaact ttggtgcttt gttctacaat acaaggagat ggaatatttt acccaggatt 8700
gcttagcagg ttacagtctt gccctctgag taccagcac ttccctgtgg gcaacatcaa 8760
cttccctgatt ttcaagtctt aattagtact ctgaagaatc ctacttgttt ttaactccca 8820
tttgccttga agtgacttta cctgattttt ttatagccct tattgcagca atgccactaa 8880
gaaactgagt ctctagcttc ttggtgggca ggagctgctt tgtgcttgc cagaatcatc 8940
cttttcagta agggagatat tgaagagaaa tctactgagg agtctggggg tgaggcactc 9000
agggaaatcc tgctccagtc caaaaagca gagaggaagg gttggttacc tagagtattt 9060
aacatgcaga gctcttggat ttactcctt taatccttgg aaatgcctat ggaaggggaa 9120
aggaagtaag atggtgactc cagcttatag acatactagt gttacatata tttaaactat 9180
aataggaggg tattattagt ttacttaac ttcaactgt gaaggattat acttctcaat 9240
atttgtctcc agtgtctatt tcagtgtatt ttcaactttt cttgaagcag catgtctgtt 9300
gcaaaacttc tagaataaat gagaatattt atataattaga tcaagccata acttgatgat 9360
atagtcattt cttcttatat tttttactta catttttaca ttttaatgat tactttcatt 9420
tttgaaaaac atgtcatgct gagatgtatt ttcttcatt ctgtaattag ttatgaaaca 9480
gtttttccta aaatgctgag tatatcaagt cttggctaag aataagtaat aaatatttgc 9540
cacatgaaag actacacata tagccaggtg cagtggcttg cactgttttt ccagctacc 9600
caggaggctg aggcaggagg attgcttgag cccagggttt ccaggctgca gtgaactatg 9660
attgtaccac tctactccag aatgggtgac agagccaggc ccatctcttc aaaacagaaa 9720
agaaagatta catagactac atatacacc ccatccaaaa catctactta 9780
acctaataatg gtaagaagat aacttcttat ttctaatat atgacacaga aaagtttttt 9840
taaaagtagt ttaaattttt aattttttct aggtatttct caagccatgt tcccatgtgg 9900
tatcttgtca acaagttgag gtggaacccc tctcagcaga tgattgggag atactggtaa 9960
agaaaaccaa ataagaacta tctcatttaa ggttaaatta cttcacataa tcaatgtctt 10020
tagctttctc taagctttat tatatattct gagtgtgttt tgaattataa gaatgaattg 10080
gggccaggca cagtactcga tgcctatagt cccagcactt tgggaggcca aggcagggtg 10140
attgcttgag tccaggagtt caagaccagg ctgggcaaca tggtgaaacc ccgtatctac 10200
taaaaataca aaaattagcc aggcattggt gtgcatgcca ttagtcccag tcacttggga 10260
ggctgaggca ggagaatcgc ttgagcccg ttgagcctgt aaagtcaaag ctgcagtga tcaggatctt 10320
gccattgtac tccagtctgg aaaacagagt gagacctgt ctcaaatata aaaagaatga 10380
attgatagag atctaagtga caacctgaca actataggta ataaaattgt attggggatt 10440
catgttaaat gagttagatt taactactct taccacaaaa acacaaaagt gggtaactgt 10500
gagatgatgt atatgttaat ttacttact atagtaacca ttatactatc tatatgtagc 10560
tcataacacc atgtcgtgta tattaaatat gcacattaaa atttgttttt taaaaaaga 10620
attgagattt tttttaacta gatatggagt ggacaaaatg taaagtgaat tgatcttttc 10680
gtctgttggt tctaggagct gcatgtggt tcccttgaa aacatcttct agatcaaat 10740
cgaatagttt ttccaaaagc catttttctt gtttgggtg atcaacaaac gtacatattt 10800
atccaaattg gtaggtgcta ttgtaattat tgctgtcata ttctacacta tagcattgag 10860
tccaaagtag aaatgaatgt gcactaatga gctttatttt ctacacagtt gcactaatat 10920

```



```

cagctgcctc ttatggaagg ctggaactg acaccaaact ccttattcag ccaaagacac 10980
gccgagccaa agagaataca ttttcaaaag ctgatgctga atataaaaaa cttcatagtt 11040
atggaagaga ccagaaagga atgatgaaag aacttcaaac caagcaactt cagtcaaaata 11100
ctgtgggaat cactgaatct aatgaaaacg agtcagagat tccagttgac tcatcatcag 11160
tagcaagttt atggactatg atagggaagca ttttttcctt tcaatctgag aagaaacaag 11220
agacatcttg gggtttaact gaaatcaatg cattcaaaaa tatgcagtca aaggttggtc 11280
ctctagacaa tattttcaga gtatgcaaat ctcaacctcc tagtatatat aacgcgtcag 11340
caacctctgt ttttcataaa cactgtgcca ttcatgtatt tccatgggac caggaatatt 11400
ttgatgtaga gcccagcttt actgtgacat atggaaagct agttaagcta ctttctccaa 11460
agcaacagca aagtaaaaca aaacaaaatg tgttatcacc tgaaaaagag aagcagatgt 11520
cagagccact agatcaaaaa aaaaattaggt cagatcataa tgaagaagat gagaaggcct 11580
gtgtgtctaca agtagtctgg aatggacttg aagaattgaa caatgccatc aaatatacca 11640
aaaatgtaga agttctccat cttgggaaag tctgggttag tataaatttt ataacttggg 11700
agaaatttta tgtggcttaa acatccccaa attatgaatt agaatagtat ttcatatata 11760
aattgaaaat caattaaaaa gaaacacagt gcctaaaggc acttggggga cacattttac 11820
ctttgcagta aagtccttgt ttggataaag attgtatgtt ttctggccaa gtaagcttga 11880
ataggtacaa gcttagatag gttcaggcca gagaggtcaa aattacttgc ctgagattgc 11940
atagctagtg ttacaactag gattcaaaccc caggcagatt gacttggggg ttcacaggaa 12000
tggagtggcc tcaaaagcct cccatcttta atgcttgcaag atttgttccc cagttaccga 12060
aagcaacttg ttaatatatt ggaaaaggcg cagtgtaggg agagatccat ggcatgaggt 12120
aaccttctcg ctgcatgtgg tggcacctgg attggaatgc atccaggagc tgcttaccct 12180
gccggtgtct gctctttaat ttgtgtataa cggagaggaa gtagacaggg caactagtgc 12240
tccagccctc catcctggcc acaaatatta atgctacctt tatatgacat aagtcactag 12300
tccatttatt ggaacctaaa tttgaaccac tgtaaagtaa gacttcatag tgataaagag 12360
aggaacttgt taggaaagag aataaaatag aaagagaagg ttgtctcctt ttgtagattt 12420
tttttttttc tccaacagtt ttacctgtga cctttataca aataactgac aaagcattaa 12480
tctctttggc ctacatcatt ttcttttcta tttttttttt ccacaagatg gagtttcact 12540
cttcttgccc aagctggagt gcagtggcat gatctggctc actgcaacct ccgctcccca 12600
cgttcaagtg gttctcctgc ctacgcctcc tgagttagctg ggactacagg catgcaccac 12660
cacgctgggc taattttttg tatttttagt agaaactggg tttcaccatg ttgaccagcc 12720
tggctgggaa ctccctgacct caggtgatct gcctgcctcg gcctcccaa gtgctgggat 12780
tagctgagat agccactgct cctggccggc ctacatcatt ttctaaagct ccagaccatt 12840
cttttctttt cttttctttt cttttctttt cttttctttt cttttctttt cttttttctc 12900
ttctcttctc ttctcttctc ttctcttctc ttctcttctc ttttctttt tttttttgag 12960
ttagaagctt gctttgttgc ccaggttgga gtgcagtggc accacctcca ctactacaa 13020
cctccacctc ccaggttcaa atgattctct tgccctcagcc ttcagagtag ctgggactac 13080
aagtgtggcg caccactcct ggctaatttt tgtattttta gtagggacga ggtttacca 13140
tggtggccag gctagtcttg aactcctggt ctcaagtgat ccgcctgcct cagtctccca 13200
aggtgctggg attacaggcg tgagccactg tgctggcctc cagatcatta ttttctgta 13260
gctttaaact gtccgttcag gagatcccac tgcatcctca aattcaaaat atctaaccat 13320
gagcttarga tttagctggg tctgtcatta gatgggaata tccttttatt tccttgaaat 13380
tatatggtga gaacaggag aagtgtctgat ggtaaagtcc tgtgattaag atagcaataa 13440
ggactccgcc cttcccactc cactgaaggt tgaagagcca tggacaatga gaagtcacag 13500
taggtgaaat caggtactaa aatggacttg gcttgagaga tcaaaattga tcacttgggt 13560
atacaactaa caaattcatg ttaacttgaa cctttattac cctgtgaagc atggtgatta 13620
aaaaaaaaca acaaacaaac aggaacttg attgttaaat tctctttaag tcagaatatg 13680
taccttagag tttttattta tgcttttgc taccattaat atgtctgcac ctgctcttta 13740
gaagttaata gagagtaaaag tcgtctttat gtctttcagt gcttacttat atttgggaag 13800
ttgagaaaaa tttttaacat cattattgat atatatatat atatatatat atatatatat 13860
atatatatat atatatatat atagataatt tttttttttt tcttgagacg gagtctcact 13920
ctgtcgccca ggccggagtg tggtgccgat ctccactcaa tgcaagctct gcctcccagg 13980
ttcaagcgat tctcttgct cagcctcccg agtagctagg atacaggctc ccaccaccac 14040
gcctggctaa tttttgtagt tttagtagag acgaggtttc accatattgg ccacgctggg 14100
ctcaaacctc tgaccttggt atccgcccac ctcgccctcc caaagtctg ggattacagg 14160
cgtgagccac tgcgcccggc tgaggtaaaa tttaaagtgt acaattcagt catttttagt 14220
atatttatac tagttgtaca gccatcacca caatctaagt ttagaacatt ttcataggg 14280
gggtgggagaa attttactct gctttttaga ttaagtttct gtctggatct aatcatttaa 14340
tcagacaatc aggcagattg tctgtgatta gttttggcca ttccagcttc ttcatgtgtt 14400

```

```

gttaactttc acaataaaag gctgctcaaa gattagaaat aacatttraat ttgaatgtaa 14460
atgtgccata gtttaaaaaga tgggttttgg gaatacagtc aaatacatac atttaaaagc 14520
ctaattctga agattatgta aagaaaagga aagaaatgta gggagaggat tgaatgttt 14580
atggtataac aatatctgaa catccatctg gtcacaccgt tggatattga atgttttgc 14640
ctcctcaaat tcatatgtcg aaatcccaac tcccaagggtg atcgtattag gaggtgtgg 14700
ctttgggaag tgattagggtc atgaagggtga agccttcatg aatgggattc gtgctcttat 14760
aaaagagaac tgtgagaaat aagtttctgt cgtttgttag ccaccagtc taggatattt 14820
tgatatagca gcctgcatgg actgagacaa ctatgagtta ttatgatagc ttctgttatt 14880
tcacctaagt tcatagaagc caatatatca atatttatgc tatgaaatat ttcttaacca 14940
agctttgaat atatttatat ttttgtttat ttttaaaattt cagattccag atgacctgag 15000
gaagagacta aatatagaaa tgcagtcctg agtcaggata actccagtcg aagttacccc 15060
taaaattcca agatctctaa agttacaacc tagagagaat ttagtgagtt caaatatata 15120
tgttacatca aaattctttt acacgttttg taagatttct agttgcttta gctaagtaat 15180
aagaatgttg tattcctttt tgatacaaat ctttttttat tgtgttaaac tatatataac 15240
ataaaatagc ccattgttcg catttttaag tgtataattc aaaggcatta attacattca 15300
taattattga caaccatcac cactatctat atccagaact ttcccatcac ccaaaagaga 15360
aacttggtac ccattaaaca ataattcccc gtccactcct ttcccagtc cctggtaatc 15420
tctaattgat attgtgtctc tatgaattta cttattctag atatttcata tataagtata 15480
agtatgcatt tgtcttatgt atctgactta ttccatttaa cataatgttt tcaaggctca 15540
tctgtgttgt atgtatcaga atgttattcc tttcatggc tgaatactat tccattgact 15600
gcatatacca catttgttta tccattcatc tgttgatgga cacttgggtt gtttccacat 15660
ttttggctgc tgtgaataat gctacagtga acattgggtg acaagtatct gtttgagtgc 15720
ctcttttcag ctcctttggg atatacctag gaattatgtt taactttttg agaagctgag 15780
aaatctttaa taaatgataa cacaataact tatatttgcc aatgcaaata tgaatatttt 15840
tggcttttaa gagattgatc attttgccac gtggttgtaa ttaaaaaaaa ttgtcccatg 15900
ttgtttcagt ataatattg tagcctaaaa gagtgcata ctgttttact ttttactcag 15960
ttaattcttt ggatactggt agagtcagga aatgagatat tgaacttaaa gatctttgca 16020
ggtgggtgct agtggctcac acctgtaac ctgacacttt gggaagctga ggtggagga 16080
ttgcttgagg ccaagagttt gagaatagcc tgggcaacat agcaagacc catctctaca 16140
aaaaaattaa aaaaaaattt aagccaggcg tggtagctca cgcctgttat cccaacactt 16200
cgggaggctg agatgggtgg atcacttgag gtcaggagtt ggagaccagc ctggccaaca 16260
tggtgaaacc catctctac taaaaatacc aaaattatcg gggcgtggtg ctaatcctgt 16320
aatctcagct actcaggagg ctgaggcagg agaaccactt gaactgagga ggtggaagt 16380
gcagttagcc tagatctcac cactgcactc cagcctgggt aacagagcga gactctattt 16440
caaaaaaagt aaaaaataaa attagacaca tgtggtggca catgcctgta gtcctagcta 16500
ctcaggagct ctagtaagt gggaggatct ctgagccca ggagttccac actgcagcta 16560
gctatgattg tgccactgca ctccagccta ggcaatatct caaaaaaaat ttttttaaat 16620
agattattag gccagacgtg gtggctcatg ccagtaatcc cagcactttg gaaggccaag 16680
gcaggcggat cacctgaggc caggagtttg agaccagcct ggccaacatg gtgaaacccc 16740
atgtctacca aaaaatacaa aattagctgc aatgtctata atcccagcta cttgggagcc 16800
tgaggcaagc gaatcgcttg aaccgggag gcagaggttg cagttagtgg agactgcgcc 16860
actgcactcc agcctggcg atacagcgag attctgtctc aaagaaaaag gaatttgttt 16920
tcctgtcttt atcgtagagg gaggaaggag agaatgggtt tggaaatggt attgagttag 16980
ccacattatg gtagatgtat cactgggcat agagaaaagg agcatttaaa acttttccgc 17040
ctaacagatg tttcttcagg ctacactgca ctattgtgc taactgtaat gtcaaatccc 17100
agacctgtgc ctatagaaca tgaacatcct tcattggatt tgtttggtca ggcttacct 17160
ttattaggaa gatcagatgt taaaataagg gtgttaaagt taagttcaga tatgaggata 17220
attcattact attccttttt ctggcagcct aaagacataa gtgaagaaga cataaaaact 17280
gtattttatt catggctaca gcagtctact accaccatgc ttcctttggt aatatcagag 17340
gaagaattta ttaagctgga aactaaagat ggtgagtaca tttgttattt tgactttttt 17400
ttctatttaa atagttgtac attttttaatt gttcttgcaa cctgtcatac ctgtgaacag 17460
tatgtgaata gtgaaatata attatgataa ttaaacagta gtttttatgt attgaaaaat 17520
atctttggcc ggggtcagtg gctcatcctt gtaatcccag cactttggga ggccgaggca 17580
ggcgatcac ttgaggccag gagttcgaga gcagcctgcc aacatggcgc aaccctatct 17640
atacaaaaaa atacaaaaat tagcctgaca tagtgggtga tgcctgtagt cccagctact 17700
tgggaggctg aggcagaagg atcacttgag ccaggagggt ctgtgttctt gccactgcac 17760
tccagcctgg gcagcagagt gagaccctgt tggggggaaa aaaaaaaaag tcttttaactt 17820
aaataaattt gacatttaaa atcttaaaatt atttcatctc tgtttcagta ctaactctgc 17880

```

```

atttattact ttttttttaa taggactgaa ggaattttct ctgagtatat ttcattcttg 17940
ggaaaaagaa aaagataaaa atatttttct gttgagtcct aatttgctgc agaagactac 18000
aatacaagta atagcatgtt attgaatat taataaaata ctatttggtt catatgattg 18060
ataataaagt atgaagtccc ttgtaacacc ttgcatgttg aagtgtatta aaaacctgct 18120
aagagtaagg aataacttga tttaaaatat tttattctgt aatctcttta aattatctgt 18180
acaaattatt gacttaacct aaatttaaaa atgaatgcct tagcacaatt aagttccaaag 18240
aatagagttg atcatgttaa ctggtaaaat gatcatgatt taaaattctt ctaggattga 18300
aacaatgaa aacgtagttt taagggtttg attttttaaa ttcttatttt tacatgcaat 18360
tttactgcac aacctatctt attttgacag ttcttaaaatt cgcaactctt cagaaatatt 18420
atcagatcac ttttctttgc ttccataagt ttttttatta ttatattatt attttttttt 18480
tttaaaagac ggtgtctcac tttgtcgccc aggctggagt gcagtggcat gatcatggct 18540
cactgcagcc tcgacctccc aggctcaggt gattctccca cctcagcctc ccaagttagt 18600
ggggaccacag gcgaatgcca tgatgcctgg ctaatttttg tatgtttgt agagataggg 18660
tttcaccatg ttgccagaa ttgtcttgaa ctctgggt caagcagttg ttctgccttg 18720
cccacccaaa gttgtgggat tacaagtgtg agccactgcg cccagctatt ctagaagat 18780
tttaagagta atcttttttt tttttttgag atggagtctc acctgtgcac ccaggctgga 18840
gtgcagtggc acactctcgg ctactgcaa cctccacctc ctgggttcaa gtgattctcc 18900
tgcctcagct tccctagtag ctaggattac aggcgcagtc caccatgccc tgctattttt 18960
tgtagtttta gtagagacga gatttcacca tgttggccag gctgctcttg aactcctgac 19020
ctcaagtgat ctgccctcct cagcctccca aagtgcctgg attctaagt taaaccacca 19080
caccagccca agagtgggtc ttttacaata ttatttttg attaggacat tcattcttgc 19140
cataaaattg aagatactct agtcatttag aatttcattg ttttggaaact agacattgtt 19200
tctttatttt tgaaatgtta ttgaaggaa accatttggg gaagatacaa atgtaagaat 19260
tgtgaaaagg ataattgtga cacaatcaa aattatagat aaaaatatac ctgtaaaatg 19320
tatttaaggca ataactctct tctgtctgt tgaccataaa tatttatatt ccctggatgg 19380
gtacattggt attgtcaagg gtgtttaaat aatgatcttg catgcataat ttattctctc 19440
tggtataaca gaatcagcaa tttagttttc tgggaccgga gaaaaacatg caaaagacat 19500
actttgaaat gtaaaactga tttttccttg caactgtagg tccttctaga tcctatggta 19560
aaagaagaaa acagtggagga aattgacttt attcttctt ttttaaagct gagctctttg 19620
gggtaagaag ttatggccaa actagcatgt tagacatgtt ttaacacta tatctggcag 19680
agttttcaat gtaaatatta aagtagatgt taatgtcaat aagtgtactt aataatgcat 19740
cagtagatat tttttcaagg attgtctcta tcttcacgcc tagcttataa tttgccttgt 19800
cgtctttttt tttttctctt tatttttatg tttttatcca tccctgggtg taggggataa 19860
cctgtctctc ttcgataaca agaagtctga agcttattag aaattttact ttgagaattg 19920
atcgatgaga agaaagcaac tagatatcac gtggatcata tatgcttgaa taaaacaata 19980
attcttagaa caaataaata cattttaaaa gttaaagcca aaaacattag ttgaatggtt 20040
aaaaatattt caaatttaagt tattccttca ctgtcttgta ttactgtaat aatttggtt 20100
atttgtgttt ttctcaactt ttaaaacaaa tatttaaaaa attcctcttt tgattaagta 20160
gggctagata aaatataaaa aatatttttt aaactcctct taatttccat atttcttata 20220
taatatgaga atctcttata aacactacct cttagaagtc tccacagaag ctttggtaga 20280
tgtagtagta gggatttgat ttcttagaat ggtataatct gtaaatgttt tagttaaagg 20340
attaaacgat aaagtcaaaa tgtttatagc acagtgttta ttaataataa ataaaatctc 20400
tttttttttt tttgagatgg actctcactt tgtcactcag gctggagtg cagtgtgcaa 20460
tctcagctca ttgcaacctc cgctcctgg gttcaagcaa tccttccgca tcagcctcct 20520
aagtagctgg gattacaagc atgcaccacc acacctgcct aattttttgt atttttagta 20580
gagatggggg ttcaccatgt tggccaggct ggtctcaagt gatccgctg cctcagcctc 20640
ccaaagtgcg gggattacag gcgtgaacca ctgtgccag cataaagtaa aatctcttca 20700
gactctcatg tgatcatgta aagtggcagg cagtcacagt caagaagtag tttaaagttc 20760
atgtttgtaa aatataatct acagattgat actggatttc ataggtaatg tttaaagaaa 20820
aataagtttt tagttatcct cagtacttca aaagcaccga tttatgatta tgttgattac 20880
taaaactaaat catttggggg cttagaggtgt ttttttatgt gttaaagattc cttaaggagt 20940
agcttgaggg caaaactttt agtaactgca tattttaaaa gtaataaaac taattttaaa 21000
cggttgaggc ctggcgcggt tggctcacac ctgtaattcc agcactttgg gaggccaaag 21060
cggttgaggc acttgaggtc aggagtttga gacgagcctg agcaacatgg tgaaccttg 21120
tctctactaa aaatacagaa attagccagg tgtgtgtgtg ggcacctgta atcccagcta 21180
ctcgggaggg taaggcagga gaattgtctg aacttgggag gcagaggttg cagtgagccg 21240
agatcatgcc actgcactcc agcctgggtg acagagcaag actccgtctc aaaaaaaaaa 21300
aaaaaaaaa gcttgaagtc agattcgaca ttaatcagta tactttctct caagtagggg 21360

```

acaatttcta	agatttttagt	cttttaaaat	ttattaacta	gtctgagcat	ggtggccttg	21420
gtctataatc	ccagcacttt	gtggggccga	ggcagatgga	tcacttgagc	ccaggagttg	21480
gagactagcc	tgggcaacat	ggcaaaaccc	cgtctctaca	acaaatgcac	acacaaaaaa	21540
cccaatcagc	tgggtgtggg	gttacactcc	tgaagtccca	gctactcggg	aggctgaggc	21600
aggaggatca	cctttgccag	ggcgtttgag	gctgcaggga	gctgggttca	caccactgcg	21660
ctccagcctg	gatgacacag	caagccccct	tctcaaaaaa	aaaaagataa	aaaattaaat	21720
taaattaatt	aactacactg	ggaaggcaaa	attcagcatt	tttttatagc	taaattttat	21780
cctgcttcag	tcttttatca	tgttaactatg	tatatTTTTT	acagaggagt	gaattcccta	21840
ggcgtatcct	ccttgagca	catcactcac	agcctcctgg	gacgcccttt	gtctcggcag	21900
ctgatgtctc	ttggtgcagg	acttaggaat	ggagctcttt	tactcacagg	aggaaaggtg	21960
agtggttaag	gtgtgttcat	ttttctgtaa	catttaataa	cttttcattt	atctttcttt	22020
gggttttgac	catctattat	atagggtggg	ttttgacct	ctattatata	gggtttatat	22080
gacatatgga	aagcattcat	ttattcacta	atatttctgt	gtgtctgctt	ttaggtgttg	22140
ggggagtgat	gacgaataag	actgatgttc	tccatgccct	ttttctgtgt	cagttgatac	22200
aatttatagg	tttttctttt	ttaggctatt	agggtgtgat	agggttgagt	aacttacaaa	22260
tgttgaacca	cctttgcata	cctgtgataa	ataccacgta	gttgtgtgtg	atcattcttt	22320
ctacattgct	gagttttatc	tgtaatgttt	ctgttgagct	tttgtccatt	taagtttgaa	22380
agtgattagt	ttgcagtttt	ctgtttttgt	gttgtctttg	tctggttttg	ctatccgtgt	22440
aaatctggcc	tcataaaatg	agatgggaag	tattctctcc	tcttcttttg	tttttttga	22500
agaggttgta	taaaattgag	gctgaatctt	gggtgttgcc	acaatgacag	gaactatttc	22560
tgtgactgaa	tatatgggga	attcctataa	agcaattatt	ttctagggaa	gtggaaaaac	22620
aacttttagcc	aaagcaatct	gtaaagaagc	atttgacaaa	ctggatgccc	atgtggagag	22680
agttgactgt	aaagctttac	gaggtatgag	tatggtaaca	ctctatataa	atcccccttt	22740
cattagaaag	acaggaatgt	tatacataat	gctgtcaatc	taataaatac	acatatcatc	22800
tagtctttaa	cttttctgtt	tatcattttag	tcattaaaaat	ttctttggct	ttctaattgt	22860
tttgataaaa	tttctaaaac	tctccatatt	taatggaggc	ctattttttt	ttctagccag	22920
aactttttgt	agactacatt	tctggaagtg	ctcactgaca	ccactctgaa	aaattagtac	22980
ttagaatata	ctctaattgg	tataaatgat	ctctgaattg	ctatggaaaa	ctgggagaa	23040
gggtgtcttca	ggggagagaa	agtaggaggc	tgtggacagc	aatgaggaga	attacagttc	23100
accatataac	actttttgtac	ttttaaaagc	cttaacattt	acattattat	ctattcaatt	23160
aaaaaatatt	gggaagattt	tactttgaac	agttaatttt	tccccatgg	gtaccgctgt	23220
catatagttc	caactaatca	tgaacttgtg	tatttctctg	tctttgtaaa	tttaaacctt	23280
gtaactcacc	aggaagtgtg	aagccaaatt	tgtgtttcaa	atatagcaac	tccaggatct	23340
ctaggcagat	gcatttgcat	ttgattttta	atgaatcttg	atcccttact	ctcacttatg	23400
ttttccacac	tcctactttt	tttattttgt	tgtaaagccat	ctaaaattct	caatgggatg	23460
aaactgggta	taaatgaata	catgcataca	ggaatttatag	tagcatattc	cttttctttt	23520
ttcttttttt	ttttttttga	gacagagtct	tgctctgtag	cccaggctgg	agtgcagtgg	23580
tgcgatctcg	gctcactata	gcctccacct	cccagggttca	agcaattctc	gtgcctcaac	23640
ctcccagagta	attgggacta	caggtgcatg	ccaccacacc	tggctaattt	ttgtattttt	23700
tagtagagat	gggtgttcac	catgttggcc	aggctgatct	caaactcctg	acctcaaagt	23760
gatctgcctg	ccttggtttc	ccaaagtgtc	gggattacta	gcataagcca	ctgcacctgg	23820
cctccttttc	tgagttttat	aaaatttgat	actttactgc	acgctttgag	actgtattaa	23880
ttgaaccatg	ttgatgaaca	agttttttgt	atgggtatat	taataaaaata	tagatcaaat	23940
ttttatagtt	aaatcaatat	cgagcttttc	tagtgctttc	aaaaggacaa	cctgaatttt	24000
cccagcactg	aaatgatact	gaaaccattt	catatcttct	gtattaagga	aaaaggcttg	24060
aaaacataca	aaaaacccta	gaggtggcct	tctcagaggc	agtgtggatg	cagccatctg	24120
ttgtcctgct	ggatgacctt	gacctcattg	ctggactgcc	tgctgtcccc	gaacatgagc	24180
acagtcctga	tgcggtgcag	agccagcgcc	ttgctcatgg	taaatgcata	caccactggc	24240
ttaaggtcct	gttcttttgt	cagtcagcat	ttttagtctt	aacaataaat	ctactctctt	24300
cagagaataa	tatatgtgtt	atgttaagtg	ttgtgtttga	ggccccctgat	ggcattctac	24360
agttgtccta	tagactgtaa	tagcaaaatt	ggtagagtaa	aaacagtgtg	aaaattctgc	24420
aacttcattg	ttagtctttt	agggtttttc	attctccctt	acttattgtt	taatttacag	24480
atttactctt	ttgttcattt	gacaaaatatt	tgtcaaatgc	ttgtgcacag	tctgtattct	24540
caaattctag	gagaaaaaga	agggtgaaca	gtattagcgc	agaacgatac	taataatgat	24600
ggctactgtg	tatgagttag	cagccctttc	ttggctttct	tggattgctt	tgtattctac	24660
atgaagatat	tccctgggct	ttacaggtca	ataaatggaa	attcagagag	attaatttga	24720
ccagggtgac	caacaaggag	atgacagcat	acactatgcg	agaagtatac	acagagtagt	24780
gtaggagcat	ataacctaaa	ctgggggtga	ggtgggataa	ggagtattca	gggaaggctt	24840

tttggaggag	ttgacaactg	agccgagttt	tgatggaaga	gtagaaatta	gcatgaacca	24900
atttcatgct	aataaagaag	caaaggaagc	gtgggtctaca	ggcaaaagca	cagaggtaca	24960
ggaagtaatg	atatgttggg	gaataccctg	ttgactggag	cttagagtgc	aaggagagga	25020
gtgctaggga	ggtgagggtg	gaggggttgg	cagcattgac	ttgcttcaag	gttcttaaga	25080
gctgaaatag	atataaaatg	caactaagag	tggcttggat	tattattacc	tagtgtgtta	25140
atctcaaatt	ttgaaatcta	tagcatctat	aggactgggtg	ttactaatct	tacactcgat	25200
ctgttactgt	tcttatacta	gatctattag	tccagtgttt	aagggagtgg	tgcatgtttc	25260
taggtcagga	caggactcag	atgtacatta	ttaatgccta	tttcagtctc	gaccttctca	25320
tatgaaacct	tataagacct	ggggtaggaa	gagattgttc	tggaagtcat	aggaatctga	25380
actgtatttt	gtttaacaaa	caatacagta	tggaaattta	tcacccttcc	agaatattta	25440
tttcagagac	aaatttttat	cattcgttca	tttatttcat	aagatccacg	agtagggaac	25500
ctcactagac	attgtctcga	gtatatggct	tgagtttgca	gtacctcttg	tgtctccatt	25560
agattttata	ggctctcaat	agataaatca	gggaataaact	agatggattc	attttctaaa	25620
gacatgaaag	agcgatacca	tacatactgc	accttaaaag	tcaaccttag	agtatcatta	25680
tttttaatag	atgtataaatt	tttaaatctc	atgtttactt	ttcctaagct	tttgactat	25740
attgcttaat	tccagctttg	aatgatatga	taaaagagtt	tatctccatg	ggaagtgtgg	25800
ttgactgat	tgccacaagt	cagtctcagc	aatctctaca	tcctttactt	gtttctgtct	25860
aaggagtcca	catatttcag	tgcgtccaac	acattcagcc	tcctaatacag	gtaatacact	25920
acttgaagg	attattgaat	tatgtccctt	ttatagaaat	tatttttcaa	ttttattagt	25980
aattcgtggc	tttaaattta	tgcttctctt	aatgatttta	aggatatgta	agtcacactt	26040
tggtgcatat	tgtgctagag	gcataaatta	taatttatag	ccacctgaaa	tgtagtagtg	26100
cgctttccaa	gaaaatgact	tttttgaaaa	tggtatttct	ttgaatgaga	aagaacagag	26160
agaaatagat	agatggcttt	taaaacacttc	attaattaaa	cttttttttt	ccaccatcac	26220
ataatggcac	ttagtccctc	ttgggaactc	atgaggggtt	tagtggtagt	gagctgaaag	26280
aaatatgttc	caggactggc	aaacatatct	taaaattctt	aaaattttca	cctagcatct	26340
accctaaata	ttcagaccct	gtgctagtta	actgctattg	aagaacaaa	gtattatata	26400
tattattaag	gataatagaa	tggtatttga	gatattggct	attgaatatg	aatatgtttt	26460
gagaaataag	ttttatagga	accaaaaaaa	aattcttaaa	ggaaccatat	attactaaaa	26520
atgcttctta	ttggagaaag	aaatgacaat	catttattaa	tgtgattttt	tcacaacttt	26580
attaagatat	aatttaagta	caacaaactc	acataaagtg	tacaatttga	tcagttttaa	26640
catatgtaga	tgccatgaaa	ccatcaccac	aattaaaggaa	acaaacattt	tcactactcc	26700
agaagctctc	tagccctttt	actacccttt	cctccctctg	tcactcccca	gacaactacc	26760
aatttgcttt	ctgtcactat	agatttgtca	acctgatttt	ctccaaatat	acattcaaaa	26820
atatacagtt	gaatacaatt	ggaaattcga	attttgtgtt	tttttcttta	ggaacaaaga	26880
tgtgaaattc	tgtgtaattg	aataaaaaat	aaattggact	gtgataataa	caagtccacc	26940
gatcttgacc	tgcagcatgt	agctaaagaa	actggcggtg	ttgtggctag	agattttaca	27000
gtacttctgg	atcgagccat	acattctcga	ctctctcgtc	agagtatatc	caccagagaa	27060
agtatgtttt	actattaaaa	cctgaacttg	gaatcttctt	tctattgttg	agaaatgtaa	27120
ttgtagtaag	acaagaatta	aatatattcc	attgtagtat	ttgaataagc	agttatttga	27180
gtagaaaatt	agtgtttcca	gctaagatga	tggcatattt	tgaaaattca	tatagtgaat	27240
ataactagta	aaagaagttt	tgtttatttt	taaaacagaat	tagttttaac	aacattggac	27300
ttccaaaagg	ctctccgcgg	atttcttctt	gcgtctttgc	gaagtgtcaa	cctgcataaa	27360
cctagagacc	tggtgtggga	caagatttgt	gggttacatg	aagttaggca	gatactcatg	27420
gatactatcc	agttacctgc	caaggtatgt	ttaaaaaaag	aaaaagtga	tacttactcc	27480
cagaagaacc	actgtattat	tggctttggc	tttatgtgtc	agcttgccca	atctccgtgt	27540
gagtcaacaa	gtgtttactg	agttaccaaa	taaatgtctt	aacactattt	taggtacttt	27600
aacaaatttt	aattttatta	atttaatttt	tattagaatt	gagacctcac	tctgtcatct	27660
aggctggagt	acactcacag	ctcactgcaa	cctcaaaact	ctgggctcaa	gcaatcctcc	27720
tgctcagcc	tccccagtag	ctagaactac	aggcatgaac	caccatgccc	ggccaactct	27780
tttaattttct	tagagacgga	gtcttgctat	gttgcccagg	cagacagatt	ttaatgtgta	27840
tgatgcagtc	tttgatgata	agaaacttat	aatggaagc	tgaggtgata	gttacagtta	27900
atacattttg	atgtataaatt	ctgtttgtct	taatcattca	aattgtagta	aagcaagatg	27960
aactgtctgc	tgggatttga	gcagaaatgg	ataggaaata	actaggaggt	agaagagtta	28020
tcaaggttca	caggactgat	gggtgaagct	agatttccag	acccgggatg	tcagtctctg	28080
aaaagcagac	ttggcagcca	tagacgaggc	agatagcagg	ataaaggaga	caaatgtaga	28140
ttgttcttca	gaagatcaga	tggtagagtc	taggaggtag	tgtgttttaa	tcagagatct	28200
gagaggcaaa	gatcattgca	tgagatcagg	gacccatgca	aaggagtggg	aaaaaaaact	28260
gggttaagga	gcctgctgca	tggcaactcc	tgggaacagt	ggccactggg	gcctgggaca	28320

tggtgattgc	agcccaggac	tgttaaaacc	agtgtgagag	aacatgggta	tggaagtact	28380
agctagcagg	atcatgaccc	cgaagctggg	atggggcatc	aagcattagt	acatggagat	29440
tcagtagatc	cagatgcagt	acatggagac	tatatgcgtc	actgctgact	ttgggcttct	23501
ttcagattgg	agcagaggta	gagggtgagtg	ggaatattct	caatagaggg	aactaaatag	28560
gcatacctaa	taaaggagac	caggatattg	cagacagtag	cctcatgttt	ggctcacctg	28620
ttcaaaaaagt	tctcttggtc	ttgagcagtg	gtgccttaaa	aggtaacttg	agaagcagtc	28680
gattatttgt	tcagcctgga	gactcttggg	atattttact	atctttgatt	gaatagattt	28740
aaatgtacac	agctctcata	acttgcccca	tgaagcatat	ccatgaaagg	cactataact	28800
gttaaaagat	tggtttgtac	tttttaaatg	tagtactttt	aataaaacag	gaaaaataga	28860
agttctgatg	cagttatatg	cattttatat	agaatgtgtt	cttaattgga	aaaaatttct	28920
cgtagtccct	ttgagttcat	ttacagtttt	tagtaggaat	tgtattttct	actgttctac	28980
ttctgtttac	taagaaaga	tggtcgtgat	taccatctga	atttttttct	tatacataga	29040
tctttagctg	ctacttagtc	atttctgttt	agacttgagc	tctttttcat	atttttttct	29100
tttgttttctc	agtatccaga	attattttgca	aacttgccca	tacgacaaaag	aacaggaaata	29160
ctgttgatag	gtccgcctgg	aacaggaaaa	accttactag	ctggggtaat	tgcacgagag	29220
agtagaatga	attttataag	tgtagaagta	tgttgctctac	ttatcttctt	tttttttcta	29280
ggtaaaatta	acataaatgc	agttagccat	ttcaaagtgt	aaattcactg	gcattttagt	29340
cattcacaaat	gctatgcaac	caccacctct	ctctaatttc	aaaacttttt	cattccactc	29400
ctcctcttgc	ttatcccctg	gcaaccattc	atctgctttt	tgtctctatg	gatttgctct	29460
ttctcgatat	ttcatataaa	acaaatcatg	caatatgtga	ccttttttgt	ctggcttctc	29520
tcacttatgt	aatgttttca	tggttcatcc	aggtagtagc	atgtatcagt	acttcatctc	29580
tttgcatgac	tgaataatgt	taccataact	tgtttatcca	cttatcagtg	gtgaacaatt	29640
gaattgtttc	taccttttga	ctattatgaa	taatgttgct	gtaaatattc	atgcacaaat	29700
ttctccacgg	atatgttttc	atttctcttg	ggtataaaat	gaggagtaga	attcttgggt	29760
cttagggtaa	ttctctaact	tttcaaagaa	ccaccaaact	gtctttcaca	ccaactgcac	29820
cattcccact	agcagtgtgg	ggggttcctg	attctccaca	tctttaccaaa	caccattatg	29880
tttctcaatt	gtgggctagt	ctcacatttg	gaaagctagt	gggagcagcg	atccatctat	29940
taaaagtgtg	atgaaattga	gtaattgagc	acctctctct	tgtagggtct	attatgtctt	30000
tgcttaaggc	aatcttcatg	cattgtgaac	agaattatac	ataaatgtct	agataaaaagg	30060
gcaaacattt	cttaaaagga	gtagacaact	agaggcagga	gaccataact	aggcaggaag	30120
ctgggggtttt	tatggtttctg	ttacttttga	ctatatctca	ccattgcttt	tgtcaaaagt	30180
agactaggtc	taagtttttt	tcaggataaa	ggtgagtggt	gtaattaaag	ggcatgtctg	30240
cagatcattt	tggttaaatgc	ttcacagtcc	accactggtg	tgtcattgtg	gtcgcagatc	30300
cagtatctta	gctgtgtaat	ttcagacatc	agcaatatta	gtttaacaaa	gggcaattag	30360
attccaagac	aaaggaatcg	tgtattattc	tagccttatt	caaacttgat	ttataaatca	30420
gtttagtatt	ttatttattt	gtttctgtat	ttatttttat	ttctttgaga	tggagtctca	30480
ctctattggc	caggctggag	tgtagtgtat	caatcttggc	ttactgcaac	ctctgcccct	30540
tggtttcaag	ctattctcct	gcctcagcct	cccagtagtc	tggtgattaca	ggctaatttt	30600
tgtattttta	gtagagatgg	ggtttcacca	tgttgccag	gctggctctg	aactcctgac	30660
ctcgagtgtc	ctgcccgcct	tggcctccca	aagttctggg	attacagacg	tgagctaccg	30720
tgcccagctc	agtttagtaa	tgtataactg	ggttttaccc	agttgtaaat	tactcttttg	30780
tcgtgttttt	ttgagaactg	gcaatgacgg	agaaactaaa	agtgccaggc	tgttgccctg	30840
ttcctgttat	tttgcccttag	tttttttttt	tttttttttt	ttctctgaga	ctgagtcttg	30900
ttgtgttacc	aggctagagt	ggagtggcat	gatctcggct	cactgcaacc	tctgcctcct	30960
gggttcaagt	gattcctgcc	tcagcctccc	gagtagctgg	gattacaggc	gcctgccacc	31020
gcaccgggtg	aatttttgta	tttttagtag	agacgggatt	ttaccatgtt	ggccaggctg	31080
gcctcgacct	cctgacctca	tgatccacca	gcttcggcct	cccaaagtgc	tgggattaca	31140
ggcgagaacc	accgtgcccg	gtcttgccct	agttatttct	tgttccctcc	tctagtccca	31200
tagttctctg	actgtattga	ggaaatgtaa	ttaaatatta	ttatgttaat	agatatttat	31260
gtggttgaat	attagaaatt	ccttattttg	gtcacatatc	ctgatcagta	gttggtcttc	31320
tggaagatagt	gattttttcac	tagagatgac	tttaggacct	attcagggtt	tttttaagat	31380
cccaattttaa	ggaaaacta	ttctcattat	tgtattttgt	atatgcaggg	aaattttatt	31440
cgaaagggtt	ttcagttggc	ttttagggaa	gattatataat	tctctttttt	tttttttggc	31500
cttttccac	atgttctaaa	aatgatataat	tctttaactc	ctatgaaaat	acattgtttc	31560
agtaattgaa	gatgctgatt	aaagtcatat	ctctacacat	tttttaaaat	ttgagataga	31620
tgggactttg	tcccttctta	caccattcac	ttattcactt	ggaaaaacta	ttatccaata	31680
cttatgtggc	agacactgtt	tctggcacaa	gggattcagc	agtgaacaaa	actgcctttt	31740
tggagtttac	attctactag	tggaaagcga	caacaagcag	atagacacat	tcagtatata	31800

attcactgtc	agatgggtggt	ggtaagtcct	atgtaggaag	aaaagcaggg	taaggaggct	31860
tggagtaact	ggagtgaagc	atagatggac	ttgtcaggaa	agggtttctg	aagagggtgt	31920
atttgggcag	agatctaaat	aaaatgaagc	aacaagccat	gagaatatcc	gggggaaaaat	31980
gttctgggca	gaagcatcaa	gcatagaact	tgtggatga	tatttattct	agcacacatt	32040
aattttaaaa	atgtataaaa	gacatccatt	taatcataat	aaagatttcc	atgattcata	32100
tagacttagt	cagaaaccaa	atttatattt	cttttttaaa	taattttatc	tcaactctta	32160
ttttacccaa	taggggccag	agttactcag	caaatacatt	ggagcaagtg	aacaagctgt	32220
tcgggatatt	tttattaggt	tggtagccta	tgaatgtttt	taaagtaact	gactctgtta	32280
ttatttatca	atcagtgtct	tttttggctt	tgttttttga	agaactgata	tttgaacctt	32340
gtggtttatg	tgaattatta	ataagctaga	ggacgtggat	tctctatttc	atcaaaataa	32400
acaaaacatt	ttagatatta	aatttttgaa	attatttggg	tttgttttac	aatagaaata	32460
ctcctcaaag	tggaatcgaa	gtggttattc	aaagaaatct	cagagtagat	tcttatatga	32520
agcaaataat	tgcccctaata	ttatctctaa	attttgttaag	ttctaaattc	ttttttcccc	32580
cagttttctaa	tttatctctt	ataagctcaag	agtcctatctg	gccaaatttaa	tttcagttag	32640
tgtaaactatt	ttgcataat	taaaaaactg	tatatgaata	cagaagatgg	tatttaagga	32700
tgaaaataat	tattcaaatg	tgatagcatt	atggggagtt	ttaaaaaaaa	agttactgtt	32760
ttattcttcc	aaaaatttta	ttataaagta	tacagttaag	agaatataca	taaaaacat	32820
atgcagctta	aggaagaata	ataaaatgaa	tacttcatgt	attcaccacc	gagtttacca	32880
ggaaaaagca	taaaacaaat	aaacctcttc	cacgtaattc	ctgggttaaa	gagaagttat	32940
agtggaataat	atttgggagc	aaacgataat	gaaaatacta	tccattaaaa	ttgttagatg	33000
ttgcaaaact	gatttcaagg	aaaatttata	gtgttaaatg	tttagaaaag	aaaaaagggt	33060
agaagttaac	cacttatgta	tctatctcat	gaaattagga	aaattataga	tataaactaa	33120
aaaatatgtt	aaaagggaat	taataaagat	aagaatgaag	tttaatgaaa	cacaaaacag	33180
agaagctcac	aaagccaaga	tttatttttt	gaacaccgag	tacaattgac	aaatctctaa	33240
caagtttgat	taagaaaaaa	gaaagcatga	ataaacaatt	ttagggataa	aaagggaac	33300
atcgctaaag	atatcccaga	aatgtaaaag	ataataaggg	aatattatga	aatattcat	33360
gccaatatcat	ttgaaaactt	aggtgacata	gacaaaaaca	aaattgacca	aaattgagca	33420
aaaaagaaat	aaaaatctgag	agttcctgta	acttagttaa	aattgagtta	gaaaagttta	33480
agaagtcttt	acacaaatca	aacatcagac	tcagttttct	aggagagttt	tgccaaacat	33540
tcaagtagca	gataattctg	gtctattttt	ggccccagaa	gatataattt	acttgccatg	33600
catttaatga	gatagtgtgt	gatttttttc	aatcaccgtg	acaggtgttt	tatataggtt	33660
gttattcgcc	agacatctag	tccacctgtt	gccagatag	gaattaatat	tcactttatt	33720
tgaattaaaa	tttgttaata	aattaataaa	acaaagtcaa	agttcaaatt	attaaaaaag	33780
taaaagaaat	aaaatatatt	ttatagagag	cccttacaaa	acagtaccaa	cataatgagc	33840
tttccaaatt	ttgaatgggc	aaaataaatg	aataggcatt	tcacaaaaga	aggaagggtg	33900
gccaaataagt	aatatattaat	ataaaaaatg	ttacttgtaa	taggaatcaa	aagtgtttga	33960
cttattgact	aagagtcagt	ttttgttttg	atccctgtta	gtctatccag	aaggcatggg	34020
tcttaataaa	caccttgacc	tcaacagttt	actgaataca	agggtaattt	catatgcctt	34080
gccttcttta	agggtttggt	gtaaaagatta	aaataaatac	ataaatatat	ataaatatcat	34140
ttatatgtat	ttatatgtaa	ttacatacaa	cttgccctct	ttaagggttt	gttgtaaaaa	34200
ttaaaagaag	tataataaata	tataataaata	cataaaataa	atacattcat	atatgtatat	34260
gaaatcactt	tgccaactat	gaagcctgat	tcaaatatga	aatgttggtt	gtttttccca	34320
gagcacaggc	tgcaaaagccc	tgcattcttt	tctttgatga	atttgaatcc	attgtctctc	34380
ggcgggggtca	tgataataca	ggagttacag	accgagtagt	taaccagttg	ctgactcagt	34440
tggatggagt	agaaggctta	cagggttaata	attataaata	cagaaataga	atgttataaac	34500
aaaatgtcat	catgtcatca	gatttttggt	aaaaaatgtt	cttttttctt	ctaggtgttt	34560
atgtattggc	tgctactagt	cgccctgact	tgattgaccc	tgccctgctt	aggcctggtc	34620
gactagataa	atgtgtatac	tgctctcttc	ctgatcagg	gacaatttca	tatttagagt	34680
ccaaaaccca	acaaatgcta	cactcttttc	ttgtgagctt	tacttctgcc	aggtaatggc	34740
aattgtcctt	agaagaccag	ctttcttagg	gaaaagcttt	agccactgtt	tgctcaaagc	34800
ataaaaagat	tctgaattag	atgcaaaagc	tttttttggc	ccagtgaag	tctgaaaact	34860
ttgtaatcct	tctgtgttgg	ctgattgggg	aaaaaaaaat	gcaagaaacc	taatgtatta	34920
tattttcaca	ttatcttctg	ttcaaaagatt	acatacttcc	attatcctgt	caaaaaaaaa	34980
actctgatac	agaatcaagc	atgtgaatcg	taagcatgta	agcagggtttc	atagagataa	35040
tttttcaact	cttctctgtc	ctgtgtgtgt	ccaactctta	ttctccaatt	tagaagcaaa	35100
caaataaatg	aatgaaagaa	cagatagaca	aatgaatagt	caaagggtata	aagtatctgt	35160
atatatgtta	catgtagcta	ttatttaaat	tatttagatt	ttccttttga	aataccttct	35220
tggcacactt	gcctaaatct	agaaaataag	cactgtgtga	ataagaaatt	atttactactg	35280

aatatattt	agggtttt	gtttttgt	ttcagaca	gtctcact	gtcaccag	35340
ctggagtac	ctggtagac	cacaactc	tcagacct	atggcccag	ctcaagca	35400
ctccccacc	cagcctccc	agtagctg	accacagg	cacgctacc	tgcccaga	35460
atatttatt	taattttt	atagagat	gggtctccc	tggtgccc	gctttctt	35520
actccaggg	tcaagtgat	ctccacct	aacctccca	agtgttgg	ttacaggcg	35580
gagccacat	gcccagcct	aagagtgt	gattttcatt	cattttccta	tatatatt	35640
ttctgttgg	gaaaaaat	caaggaag	aaatagtag	ctgttggt	atttctca	35700
ttacttata	agcttttt	atatataa	ttatattat	aagaaaat	taagatac	35760
aatttaaga	aatatttt	attttatt	ttatttgta	aataaatt	tctccttt	35820
ggtgtcac	cttgaaatt	taaatgtct	cagtgaact	ctacctct	cagatgat	35880
tgaccttc	catgtagca	cagtaact	ctccttact	ggagctgat	tgaaagct	35940
actttaca	gcccaatt	aggcctta	tggaatgt	ctctcgag	gactccag	36000
aagtatat	aggaaagt	tatgacatt	tatgagtga	aaaagaag	caatgtcaa	36060
atttccacc	taaaaaat	tatttttta	acaacttt	taaaactgt	tagaaacata	36120
aatttacct	tagtgaaat	ttccatagt	ggaatatgg	ttttgcag	aatttata	36180
tatgaagtt	gatgtctgt	tctttaac	taccttaata	ttggcaaaa	catgttgg	36240
tttgcaagg	tattattta	attgggata	catgaatt	atactacaa	caaaaata	36300
tagagtttt	tgtttgttt	tactttaact	tttaaaaa	aatcagtta	agttgttgt	36360
ttgaagctc	cattgttcca	atctggcca	taggagccc	ttttgtatg	ctcctgtat	36420
tttatgac	gtcctcat	ttcttgaat	acttccct	ttccagata	agtaagtt	36480
tcttggcc	gtgcagtgt	tcacgcct	aatcccag	ctttggcag	ccaaggcag	36540
aggatcatt	gggcctagt	tgagaccaa	tcatgggtg	acaaactgt	cccactatg	36600
acaacagag	gggatctgt	ctctgtgaa	aatttaaaa	ttagctggg	atgggtggc	36660
atacctgt	tcttagctt	ttgggagag	ctgtggcag	aggatcgct	gagtaaat	36720
aggatgcag	gagccatgt	tgtgccact	cactccag	tggaatgac	aatgagacc	36780
tgcccccaa	aaagaaaa	attcttgg	tatcttgt	tttctgtat	ccagccctag	36840
catagcctt	ttctctaa	acagtatt	gattttaata	tttacagta	atatattga	36900
tgttacatta	tagacttt	catataatt	ctaggaagg	ttattctatt	actctcttt	36960
accacattt	tttggaat	ctacagaac	tacagttct	aaatcagaa	ctccctagg	37020
ttttgctatt	ttggcaagg	attgaagtt	ttccctctc	ctttactacc	agaaagggt	37080
gtattttag	agctctct	aatgagaa	cactctata	catggttg	tcactattt	37140
ggagttaga	agtatgaat	gaaagtcag	gacataaaa	taaaagccc	aggctctg	37200
cttagcttca	ttacagact	tcttggggg	tggttggt	attatctac	cattctat	37260
tgtctttata	attttaat	ttaaatttt	accatgtgc	tcaaaacc	tagagaatta	37320
atgagctct	tgaaaaat	ttctaagtt	cttgtattg	tctaataga	tgctatct	37380
gttattatt	atttctgag	ctaaaatt	ttacatctt	aaactgggt	tccttttgt	37440
tatttttag	tggaagtt	agctctgata	gtgacctaa	tctgtcttca	atggctttt	37500
tttaaccat	cagtggct	gacgattcag	ctggagatg	agaatgtgg	ttagatcag	37560
cccttgttt	tttagagat	tccgagatc	ttccagatg	atcaaaatt	aatatgtac	37620
ggctctact	tggaagct	tatgaatcag	aacttgaaa	tggaacct	tctgatttg	37680
tatcttgtg	agtcattc	atacagttc	gaaatataa	gctatatgt	gggtgaaag	37740
tgcaagtgt	tctctccta	ccagccccc	atattcttc	tggttggtt	gttcttcag	37800
aaaatagtc	tgtttcttg	ttactacta	tggtaattt	cattccttg	taagatttt	37860
aagacaggg	tgggagcaa	gaaccaaag	agcgcgtgt	tgtgattacc	tttggtttc	37920
ttgaggttt	tcttaccta	tggttttaa	acatcttt	gagcagttc	attttatag	37980
aaacttaaa	tctgttat	tgaacagtt	aggataatg	ataatttg	acaataatg	38040
aagaaattc	tgaaaaaaa	gtgttatct	tgatacttt	gctgcata	aagcacaat	38100
aagtgtact	ataatgttt	aacaggaaa	tgttttgatt	aaatgtggg	agtatcact	38160
ttctactag	attcaacat	tcttctaaa	attaatagt	gttactgt	attttattg	38220
tacatgta	atctgtac	gtgtttgtt	atctatatg	ttcctgggt	tttgtacat	38280
tgctttatta	atttaggct	ttttttttt	tttttttga	gacagtctc	ctctatcat	38340
cagctagag	tgcaaggca	caattatgg	tcactgcag	cttgacctc	tgggcttag	38400
tgattcttc	acctcagct	cctgagtag	tgggactac	ggcacatgc	accatgcca	38460
gctaatttt	gtatgtttt	tagagacgag	gtttcaccat	attgcccag	ctggtctca	38520
actcctggg	tcaagctat	tgogtgcct	gacctccca	agtgtctag	ttacaggtg	38580
gagccactat	gcctagccta	actcagact	taaaaaata	aaagcaatt	attttattc	38640
ccaagaacag	taaggtggt	gtttaattt	agtcttta	tctgttttt	attttattc	38700
tttagaaatg	tcccagaa	ttagtata	tttacttt	gaaaatga	aaacctgt	38760



ttgggcatta	gtgtgttggg	tttaagcaac	aaagttaaaa	aaacctaccc	tgtgttatgg	38820
caattttcac	ttgatgggtg	ttctataaca	caggatcag	tgaaccttta	taaaagatga	38880
acaacttttc	agcttgctta	atttcagtta	attaacatgt	atacttatct	atgttaatgt	38940
tttattgctt	aaaatgttta	atttttatat	ttggtaaaaca	gatagttttt	tctctcccc	39000
tcttccttcc	atctttcatt	actacaattt	accatgcaga	gctcacaatg	tctctctgca	39060
ccaagctcca	tgactcagga	tttgcctgga	gttcctggga	aagaccagtt	gttttcacag	39120
cctccagtg	taaggacagc	ttcacaaag	ggttgccaa	aacttacaca	agaacaaaga	39180
gatcaactga	gggcagatat	cagtattatc	aaaggcagat	accggagcca	aagtggagta	39240
tggtcttttc	cccctcatta	taattgttaa	aacttcttaa	aaattgtttc	acccttttga	39300
tatatatttc	tttgacttat	aaacgagcta	tatttataaa	caagggacca	gaacacatta	39360
actcagtc	ggttatgtgc	ttccttgctt	tcaatgtttc	attatcttat	aagggaagaga	39420
acgtatgg	tcttgaaaa	actgacaata	agaagtaaca	actggactac	cacatttttt	39480
tttcatcct	taatttaact	cttcgtcaat	ttcttttttt	acttaaggag	gacgaatcca	39540
tgaaccaacc	aggaccaatc	aaaaccagac	tggtctattg	tcagtcacat	ttaatgactg	39600
cacttggtca	cacaagacca	tccattagtg	aagatgactg	gaagaatttt	gctgagctgt	39660
aagtaacaga	ttctgttttg	gaagtacagc	tactattaca	agtgcacatg	tattacactt	39720
aaacctttta	agttcgtgtt	taaaataaaa	atattttgaa	tatttataaa	ctaattcaaa	39780
aaatatgtgt	cgtagctatg	cattaaaaaa	ccccaaaatg	tcagaagtac	agaagtcaaa	39840
attgagtttt	cattaaccag	ttcatttgat	tatatattgaa	ttattcataa	tggactcatt	39900
taattttagt	aactttgggc	tgggtgctgt	ggctcatgcc	tgtaatccca	gctctttggg	39960
aggccaaggc	aggtggatca	cctgaggtca	ggagttcgag	gcaagcctaa	ccaacacggg	40020
gaaaccccat	ctctactaaa	aatacaaaaa	ttagccaggt	gtggtggcat	gtgcctgtag	40080
tcccagctac	ttgggagggt	gagacaggag	aattgcttga	accaggagg	tggagggtgc	40140
agtgaagcga	gattgcacca	ctgcactcca	tccagcctgg	gccacagagc	gagactgtgt	40200
ctcaaaaaaa	aaaaaaaaaa	atttagtaac	ttcgaagaaa	taagaaggaa	aattaaagt	40260
tgaaagtgat	tctaattgat	agtttataaa	attttggtat	aaaaatacct	gttttgctt	40320
caaaataatt	tatatataaa	ttttattgac	ctcaagaaca	tttaaatata	ttcagattta	40380
ttcattttgt	gaccacattt	gttatataat	ggattttaag	gatccttgca	attgagttta	40440
tggccaccta	ttgcatctgag	accatggac	tgggaacat	tctaggtcaa	tgattcagtg	40500
tgattcaatt	taagagatgt	ttattcctgg	tctttagaag	ctgtacacct	ttgttatcta	40560
attttgca	actttgaagt	atgtatgtat	gtgtacatac	gttagtgcta	tgtattttat	40620
aaagaagaat	cagaaaacag	aggtaaggaa	aaataaggaa	acaaatttct	gttaagccca	40680
ccacctccca	aagcatattt	gtttatatgc	ttatatatgt	tttctatta	tggttaagaac	40740
agtctgtaca	tattgctata	tagcagtcct	cctttatcca	catatctcct	gaaaattgtt	40800
ttacatttta	aatgttaact	actttattgt	ttttaaatgt	cattttatag	tgtagctatg	40860
ccacaatatc	caatttttag	acattttaat	tgctcccagg	caatgtggta	atgaacattc	40920
ttgcagctga	atatatgcac	atattctaatt	gtttcactag	gatagaggtg	gaattgtata	40980
acagggagct	cacatttttt	aaggcttttg	aaatgtattg	ccaaattgcc	tgccagatat	41040
actgcaccat	cactaacatt	gtgtgttgca	gtatttttct	aaacttgccc	cttttgattt	41100
tagaaaaatg	atatcaataa	tttacatttc	tttgattaaa	gtgtagaagt	tataattttt	41160
catattattc	attgtcattt	gtattttatc	ttttctaact	tgtctcttca	tcccctttgc	41220
tccgttttct	attggagtgc	aactttattt	gtaagaattc	tttttaattt	ctgtgactgg	41280
aatttttttt	tctagtttgt	tatttcccgt	tcatcttcta	aaatataatt	gtgtttgcca	41340
acaatccatt	atcttttgtt	ttgtaatggt	agtatttata	catatttaaa	tatctctttc	41400
tttttttcaga	tatgaaagct	ttcaaaatcc	aaagaggaga	aaaaatcaaa	gtggaacaat	41460
gtttcgacct	ggacagaaag	taacttttagc	ataaaatata	cttctttttg	atttggttct	41520
gttaagtgtt	ttgatggctt	ttccatattg	tgtaaacagga	aaaaaatggt	gtctatgaat	41580
ttcttcttaa	tttaacaaat	ttggtttaatt	tataaaatca	cagattggta	aatgctataa	41640
ttatgtaatg	atcaggattg	agattaatac	tgtagtataa	attgggacat	tataacagat	41700
tccatatttt	atttcctaaa	atctaaattc	agtctttta	gaaataatat	tagccaaatg	41760
gtggaactaa	tttatttctt	ttgaggaaaa	gataataaag	aatgttaata	aatttaaat	41820
tcttgggaatt	cccagttgta	tattcatcac	ctttgtagca	tttgacaaa	tttatgctta	41880
gcagcttctt	cactgttttg	aaataaaata	tcctattacc	tactgatata	attatctgtt	41940
ctttgtatat	caaaaaatgt	gaaatttaca	cataattcaa	atacatttaa	ttatccgctc	42000
aaccagaaat	gaaatcacat	ccctctacta	tactacatcc	agctccaagc	ccaagatatt	42060
taaatgacat	ccattcctct	cctagtcca	gttatgattt	tatcttgata	ttctctcata	42120
tatgaactaa	attataaagt	tagccaccat	caatacaatc	tgcgtatcta	atatcttaac	42180
tatatagtaa	tggggtaagg	gaacagcaaa	aaggagaaca	ttaattaaaa	tatacaagta	42240

```

agcctgggca acatagtgag accccatctc ttaaaaaaaa aattagccat gcatgatggt 42300
atgcctctag tcccagctac ttgggaggct gaggtaggag gatcacttgc tcccaggagg 42360
ttcaagggtc taaaccagca aagctcagaa tcccaggagg tagaaacaaa gacttagtgg 42420
atcactagta ttaaaactgag acacgtcacc ctgcattgca ctttgtttct cagttccttg 42480
atgaaatcac tgagctgaca tacctgccct cttttcacca taaagtgagt ttcatgatca 42540
gaagcaatgt ctatgggata gcctaacaaa caatgtaaaa accatttagt aagttcatga 42600
aggggtggtg tggtaaaaaat ttggagaaca taaaaaacia atacaattcc aaggtgtgtc 42660
ccctccagga aggacaaaatt gctgcctgct ctgtgataga agaggatcag atgtaatcaa 42720
cctgccgtca gacttgggct gttctctcct ggggtgtggac ttgcctggtt ggtcactgct 42780
gctgacaagt aggtctgtcaa tatagctggg ttgtcatgtc agctgtggtg agggggaagt 42840
ccacattgtg gaggccacat ccctgcactc ttggccaatt tgaccatgaa tcttaagcac 42900
tgggggtggct ggaaaagaca gccgattgac atccatacag aggtcatctt gaccacttga 42960
ttagtataag cactgaaggc ttttaactga gcattcacat aggacacaaa tattctgatt 43020
ctttggggccc attccaagaa ctctgggcat acttttctct cagacctcat acccagttgt 43080
gttctttcca aatttctggt catctggtta tgttattagc cactatctgt gaatcagcat 43140
agatctttat actagacatc tctacctcct gacagaatgg aggagatag ttacttaaca 43200
attctgttcc ctggaagat ttctgtctc cactgtttgt aagggtact ccctcaatgt 43260
agcagttaag ctttcaactc gatgggaagt cacagtggaa ttctgggtct ccaagaatta 43320
gtgttagtgc atacacagtg tctgataatc cccagagtgt ctggtgccct tggatcctgt 43380
gaagaaggct tggagaaaag aagattcatg gcaagaactt gtgatgtgat gacagggcc 43440
tttctctggc tcttcattct tagctgacc taggtgtgag aattaggcca ggggccatga 43500
ctatattgtg gtgactcaaa ccaggccttt gtttactaac tgggagattt ttacattgta 43560
aagtcaagt aggatctttg cccatgtatt ttggtcttaa gaacacaaat gatattgctc 43620
caaagtctgg aggaacacca gggctcttgg tctcacgctg atttagataa aacgactgtc 43680
aggcctctga gcccaagcta agccatcctc ccctgtgacc tgcacgtata catccagatg 43740
gcctgaagta accaaagaat cacaaaagca gtgaaaatgg cctgttcctg ccttaactga 43800
tgacattcca ccattgtgat ttgttctctg cccatcttaa ctgagcgatt aaccttgtga 43860
aattccttct cctggctcaa aacctcccc actgagcacc ttgtgacccc cgccccctgc 43920
cctaagagaa aacccccctt gattataatt ttccactacc caccacaaat ctataaaatg 43980
gccccacccc tatctccctt cgtcgactcc ttttctggac tcagcccgcc tgcacccagg 44040
tgaaataaac agccttgttg ctcacacaaa gcctgtttgg tggactctct tcacacggac 44100

```

&lt;210&gt; 64

&lt;211&gt; 16869

&lt;212&gt; ADN

&lt;213&gt; Homo sapiens

&lt;400&gt; 64

```

aagctttagt agagatctca aaaatggttg gatggtagca aattactaag aactctcaaa 60
gtttctaaag ccttagtttc agcttgctag aaaacctatg ttgagtatta tggctagttc 120
catagttgag ttgggaaatg tctttgagga gacacttttt cactttgtat tcatctgtac 180
atcttctggt acttgcatte tgctatgtc aggtatttag agcaggtaca tttttataac 240
tggaatgttt atgtgtagtg aagctctgag aggactttgc attagatctc agcagcataa 300
tcagaagggt gtcctttgtc tcagcaattt ttaagctaag agtagcagaa attgcagtgg 360
aaatagactg ctttgccaca acattcagaa aatcatttat ctttttattg cagttcttgt 420
caccaaacia tacattttag tacttctcaa attgcagaac tctcataggg ctgggaaaaa 480
gcctgtagac acatacatat tatgaatgtg ctaatgtttt ttgtattttc atagccatc 540
aaagctcctg agtcagtttc cactataatc actgcagaat caatcttcta caaggtaagc 600
ttttgtagag ttactgaagg aagagttggg cctagtgggt aatgtgccac taaaatgttg 660
gattagtcta aaggtctctg ctactcttta tttgtataag gtgtgattat actttttgtt 720
cccttcttag ctgttttccc ccataagtg ctgttattaa aacatctcat cttagagctga 780
agtgggagga gaaagtgcct actgacacat gatgtgagga tcttaagtat ttttttttag 840
tgtagattgt aggaattatt cttaaaatgc tgattgtata gtgtggagcc atggaagact 900
gagccgttag tgcatgggca ttgaagaatg agaaggacag agacaggatt tggactagta 960
gaggttgtcg actgtgtgtg caaatgggta gagtaggccc agagattcta aaatgccttt 1020
aagtggagtt gagctgagta agggcagtag tgaggattaa cacctactag aaattcatag 1080
tgagaggaat tccaagatgt tttgataaaa gaatgaggag gtcagggttc ccagggccaa 1140

```

agtccatgaa	catctgatac	ctcagtgaga	gaagtgcag	attgttgtgt	ttaaaccaga	1200
agtcttagga	aaggaattag	aacatagacc	cccaaggctc	ggcaggcctg	gcacggcaca	1260
ggcagcaacc	attgaaggct	atttggtgtt	tgggagctct	aactgtcatt	taggggacag	1320
tgggtgtgag	tagtacttta	tacttgaccc	agggtgactg	agaaactcaa	gtgatgatgc	1380
ccttaagtat	actttttttt	aagccacaaa	tctatatagt	cgaagtctgt	tcctcccaac	1440
aggggtacac	tggcatttct	cagcagggct	gggaaaaacc	aacaacaaaa	aaagtctgta	1500
cacaggcaaa	catctctctt	atttttccaa	catcttaatac	attgttaata	aaatatctaa	1560
agttttagca	acagttgtgt	tgtatcagtg	gctgagcatt	ttgcatgctt	tatttcattc	1620
agttcactct	atgaggtgga	tactactatc	cccattttct	agatgagaac	attgaggcac	1680
agcgagggtta	attaacttgt	ccaagatcac	atagccaaca	agtcatggag	tgaggcagtc	1740
tcatgccaga	gcttaagcct	agagcatagt	tcctggctct	acagctttag	caagtgcagc	1800
gctatgtgac	gaggaccaac	ctctctaatt	tctcatctgt	aaaataggaa	ttgttaatat	1860
ttactacctc	agtgggtcaa	atgaaatcat	atgtgttaag	cacttagcag	agtaagcact	1920
caatgaatag	taggagttat	cacatcttcg	tatttgtgca	ttaccttcac	agtttacaga	1980
ttaaggccag	aagcaacttg	ttgagctacg	ggtttagtgt	actaacagtt	tccatgtgtg	2040
tctccatggg	agggtgtgtg	ggacctgtta	ttgtgactgt	ctgtactttc	gtattgttgt	2100
ctgccaccca	tgtttattaa	atgataagga	caataatgca	acaaagtagt	caagtaatgt	2160
tgcaaatgcc	cagtattgta	gtggctatca	cagcagtgcc	actggcaggc	agcaccatgg	2220
tggcaagttc	aagaggtcac	tgccagccac	tgagctagag	cccagatcag	gcatgcaaga	2280
ggagcctgag	tgaggccac	tggggatcac	ggccaagagt	gtgaccacc	aagaccaga	2340
atggctgagt	ggcctccctg	gagcatggca	gtggcagaac	aactccatga	actcagatct	2400
ggtgatgcct	aaactagtcg	tgttctcgtg	tggaccocct	ctctctacca	gaaaccttga	2460
atcctctcag	caaatgagga	gactactcag	atcagtgcct	tagtctgtgt	tgggtgtata	2520
tatgtgtaca	caacacagca	catatttaata	aatacctact	atgtgccagg	cactgcctac	2580
cactggaatc	tttactaag	acattgtttt	tactttgcat	ttctgccttt	acactatgaa	2640
agtagatgtt	ttggattcat	attcattcag	catacatttg	aatatgctgt	gttatgcata	2700
gtaagcctat	gataagcaag	tattctcatt	tagaatttgg	gaatattgat	tatacatgtg	2760
gacaaacaaa	ccataaatgc	aaactattta	tatgataaat	aactttggac	tgatggctgg	2820
gagggaaggac	cagctattga	tgggtaggaa	ctagcaagta	gaggactgtg	gcctgcatag	2880
accagaccca	tccgtagtga	tccagatgaa	acagccaccc	tcagacactt	ggataaaggg	2940
tccaccagga	aaaaactcct	ggcctatcag	gtgctatgtt	acagttcagt	tactggaaag	3000
atctcctcaa	aagtgttttt	atggttgagg	tacacattcc	tacagcttta	cctgctgcca	3060
agtcctctgt	tcaagggaag	cagcaatgaa	ttacactggt	cccgtagtca	aggacagtat	3120
atcttaccaa	gaactatacc	cacttaagga	ggtgctggat	gtcataaaga	tttggatcaa	3180
ccattatggg	tgttcagagg	agagattatt	tccagctcaa	gacccaggga	agaggacata	3240
ggatggatac	cagagtcata	gggaggattt	aacacaggac	atgtacacat	tagttagtgt	3300
ggatataaag	ggaacagaaa	tgaatgagac	acaaagcctt	gaatgccaga	aatactagta	3360
gtcctgttgt	ggaaggatat	aaaactcaac	tgggagtggg	agagaaaagg	agcagtgcag	3420
ctaggagatg	tacagttagt	tgaaggtaac	atatcctgaa	gactataatc	caaagattat	3480
ttttggtttg	aatttgtttt	ggtttgaatt	catggatatc	atcttctttg	agtggatggg	3540
tggggagggt	ggcatgtaga	atgcattcct	accaaatac	catgattttc	aagacagtac	3600
agagaaaaga	ctgctgagct	gatgtaggag	ctttggctgc	agtctctatg	gctttcagca	3660
agccgtttaa	ccttactact	gcttcatgac	tgtggctaac	aaagtaggga	tagtacggag	3720
cacagaggat	ttttaggggc	gtgaaactat	taatactctc	tttgtatgat	actataatgg	3780
tgggtacatg	tcattataca	tttgcccaac	cccacagaat	acacagcacc	aagagtgaac	3840
cctaattgtg	actctggtct	ttgatgatgc	tatgtcagtg	tacgttcac	cgtgtaacaa	3900
gtgtaccact	ctagtgggtg	gaggggttat	tgataatagg	ggaggatgtg	catgtgtggg	3960
ggcagggaag	atatgggaaa	tctctctact	tctgctcaat	tttgctgtaa	acctaaaaac	4020
tctgtataaaa	ataaagtcta	ttttttaaaa	agtggggatg	gtattacggc	aatataaaat	4080
caaaatactt	tatgaacaaa	tcttttctcc	agatgtaaac	tgtcatatat	gcaccctcgt	4140
atgtgtatgt	ataattttca	ttcaaactgt	aaacaacttt	agaattggca	ccaaacatat	4200
aaactactgt	acattagact	atctcgaaca	ccttttactg	accactttga	aaacttgctt	4260
acctattaa	gttcattcat	agctgtgatg	ttctattttt	atcttcaatg	tgggattatc	4320
ttctgtttcc	cccaggaggt	atattaccaa	attggtgatg	ttgtttctgt	gattgatgaa	4380
caagatggaa	agccctacta	tgctcaaatc	agaggtttta	tccaggacca	gtattgcgag	4440
aagagtgcag	cactgacgtg	gctcattcct	accctctcta	gccccagaga	ccaatttgat	4500
ccgcctcct	atatcatagg	taagtttgac	aaatggcaca	ggtttttttt	taacttagtt	4560
aactctccaa	tattatgtaa	aagagtgtgt	tagtcagctt	gggctgtcag	gacaaaaat	4620

cacagactga	gtggccttaa	caacagaaa	gcactttctc	acagttgtgg	aggctgaact	4680
ccaacatcaa	ggtgctggca	acacggattt	ctggggaggc	tttcttcct	ggcatataga	4740
tggtcacctt	cttgctgtgt	cctcacatgg	ccttcatggt	agtggagagct	ctttggtgta	4800
tcttcttata	aggacacccat	ttctgtcaga	tgaggggccc	acccttatgg	tttcatttta	4860
ccttaattgc	ctccctaaag	gtctcatctc	caagtaccat	cacattgggg	attagggctt	4920
caacatataa	atttggaggg	tggggggggg	ggatgcaatt	cagtcataaa	caaaaaaagc	4980
atgagtatta	ttaagtacaa	aaaaattaga	gagctttata	gaaaaataga	ggcattttta	5040
gtagctggag	tgtgagtgtt	atcagttatt	ttgagttaga	gcaatgtgca	tctactaaga	5100
agtggatagg	ataagatttt	ttggagtga	cccagggtta	aactgtacta	caagaatgta	5160
ttgctcagga	actaggttat	ttaggttact	tatttatata	aacctattca	aaaataattt	5220
aggaaagaac	tatccaggtt	atccataact	tgcaaatctt	caatatgtgt	gcctctgcat	5280
gctacacatg	tcatcttagg	cctttatagt	ataaaggctg	atagttgaaa	tggcagctgc	5340
tgtgcttttg	ttaatttcaa	agctgccaaa	acagttgtga	gatagactca	caagaattta	5400
ctgattaata	caatttttaa	agtttttcaga	tttttacagt	tacttcagac	tttttatctt	5460
tctgcagtga	gcatgcatca	ttacttttgc	atcctgagaa	caagcataag	tgtgtttttg	5520
gagagaactc	cagggacaaa	taatatacca	ctgttattct	cacctatatg	tcaagtttga	5580
tacattacca	aacaattcta	gccttctgct	tataagtata	tagaattttt	atttacctta	5640
tctatggatc	aggatctcag	cagaggcagt	gatgtatcag	aatcaccttc	gggattcctc	5700
tactgcctcc	tctttcta	ccccagattc	tgatatgcat	ccttgcctca	cagcgaggca	5760
gcatggcatg	aggtcagaac	accagttctg	gagccagact	gtctagggtt	acagcctgcc	5820
atttaccggc	catgtgactt	tggcaagttt	cttagtctct	cttgccctcac	tttccctata	5880
tgtaaaatgg	gaataataat	agtgcctacc	tcagaagggt	gatgtgagga	atgaagggtat	5940
tgtacatgtg	aaacttagag	cagtgtgggt	acaaaataaa	catgatgcaa	gtgttcaatc	6000
actgtttttg	ggagaatgcc	atattcttta	agccgttaaa	gaagaaaaaa	tgattaagaa	6060
taattttcaa	gtaatgcatg	tttcaagggt	taatgccagg	ttgtccccag	agtggctctt	6120
cccagtgctt	agaaatttta	acatcttatg	aaaatgatat	atatgggtcaa	aaatgtattt	6180
aaactttccc	ttggctgcct	tccagggccca	gaggaagatc	ttccaaggaa	gatggaatgc	6240
ttggaatttg	tttgtcatgc	accttctgag	tatttcaagt	cacggtcatc	acctttccc	6300
acagttccca	ccagaccaga	gaagggttac	atatggactc	atgtggggcc	tactcctgca	6360
ataacaatta	aggaatcagt	tgccaacctat	ttgtagtcca	caaatataaa	ctgggtttcc	6420
aggcctgggt	tggtggctca	cgctgttagc	cccagctatt	gcaccactgc	tctccaaagt	6480
gggcaatgga	gtcagattct	ctttcttaaa	aaaccacaaa	aaaactggat	ttccagttct	6540
ctaataattct	tagtaccaca	agatatgtca	taggtatctt	taaatgaaat	tcttagctgg	6600
aaaagtgact	aaaaagtttt	tctcctgcta	cctagtaata	aacaaatcat	tgtttattac	6660
tggtcactta	gaaaattaaa	agggataggg	ccaggcacag	tggtttatgc	ctgtaattgc	6720
agcactttta	gaggccgagg	caggcgatc	acctgaggtc	gggaagtgga	tcgcctgagg	6780
tcaggaggtc	gagaccagcc	tggccaacat	ggcgaaaccc	cgctgctact	aaaaatacaa	6840
aaattagcca	ggtgtgggtg	catgtgcctg	taatcccagc	tatttgggag	gctgaggcag	6900
gagaatcgcc	taaaccaggg	agggtggagg	tgtagttagc	caagattgca	ccgctgtgct	6960
ccagcctggg	caacagagtg	agactcttgt	ctcggaaaaa	aaaaaaaaaa	aaaaaggctg	7020
ggcacagtgg	ctcacgcctt	taatcccagc	actttgggag	gctgaggcag	atggatcgcc	7080
tgagggttgg	agttcgagac	cagcctggcc	agcatggtga	aacctgtctt	ctactaaaaa	7140
tacaaaaatt	agccaggtgt	ggtggcgcat	acctgtagtc	ccagctactc	gggaggctga	7200
ggcaggagaa	ttggttgaac	ccaggaggcg	gaggttgtag	tgagcagaga	tcgtgccact	7260
gcaactccagc	ctgggtggac	agagcaagac	tccgtctcaa	agaaacaaac	aaaaaattaa	7320
aaaggataga	atataatgaa	atatattttg	aacttaaaat	atattctata	tgtgtatctt	7380
cctaggcaaa	agctgtaatt	tccagagaga	ccattaggaa	caggtagtat	ctatttttct	7440
ccattatttta	tttctagaaa	ctcataaaat	ggattgtatt	tttctataag	aacaaaaatat	7500
taattaaggt	atagatgact	gaccaagggt	ttaatcaaat	aaaatgacta	acagcatcta	7560
tcataaagcc	acacaagcct	tatgttctca	tctcaaaaat	gctgtgacag	ctttttggct	7620
gctttaacca	taagaaaaat	gattgggtga	tgattttatt	agcccaggct	tttaaaaact	7680
ttcatctagg	ccacgtgcgg	tggctcatgc	ctgtaatccc	ggcacttttg	gaggcctgag	7740
tggtatggatc	acttgagggt	aggagttcag	gaccagcctg	gccaacatga	tgaaaccctg	7800
tctctactaa	atatacaaaa	attagttggg	tgttatgtgt	catgcctgta	atcccagcta	7860
ctcggggaggc	tgaggcagga	gaattgcttg	aactcgggag	gtggagattg	cagtaaggcg	7920
agatcgtgct	actgcactcc	agcctgggtg	atagagcaag	actgtctcaa	aaaagaaaaa	7980
aaagaaaaaa	ttttaattta	atccttctgt	agaaacaggc	attcagaacc	attccattga	8040
tcttaataaa	gctgctcttt	actgtttcta	gtcaaaaaat	agacttcgat	caaaccataa	8100



ctgtgggcct	ccacaggttt	gctcacagca	aatgggtccgt	gacagaaaga	cgcaagggca	11640
gttgcaccca	agatggaagc	caccatcttt	tctataacct	aatctgaaag	aagggaacata	11700
ccagcacttc	tgccatatgc	tgttgggtca	cacagacca	ctctgggtaca	gtgtgaacac	11760
aggaccacac	aagggcgtga	attccaaggg	cagagaccac	tagggaccac	ctcagaggca	11820
cagagggaca	ccctatccag	ctggtggcca	atgtaaatta	acatagcttt	ttagaatagc	11880
aatatgtatc	tataatctta	aaagtattaa	aagtacttct	tgatccagta	atttcatttc	11940
taagaatcca	tgctaagagg	atttaaaatg	tggacaaaa	aatgggtata	aaaagaagtt	12000
gttaacagta	tttaaaagttg	tgaaaaacca	gaaacaatct	aaagggtccaa	caataggaaa	12060
atgaattttg	atatttttct	aatagaattt	tatgctgtca	tcagaaatac	catttacaaa	12120
taatttttta	taacgcaaaa	aaaagtttat	aaaatgttta	gtgtaaaacc	tggacacaa	12180
tacataatga	ttctgatttt	gtaaaaaaa	aaaaaaaaaa	cacacacata	tacacatgca	12240
tacatatgca	tataaagaaa	actggaacaa	acaaaataac	aagcatagtt	ggaattacag	12300
tcattttaat	attctttatg	cttttaaaaa	ttttgaagtt	tgtattacta	gcattccacta	12360
cttacgtagt	caggaaaaaa	atacaacttt	aaaatagata	tttaggtcca	aagatggtaa	12420
tctaaatggg	gttacaggct	gaatgtgtgc	ctgatcccca	tgccccaagt	tcatatgtta	12480
aagccctggc	ccccaaaggca	atgggtattag	gggagtaggg	cctttggggag	gtaatcagga	12540
ttctacgagg	tcacaggggt	ggagcccgca	tagtggaatt	agtgtccttt	taggaagagg	12600
agaacagacc	aaagccttcc	tttctctcct	cactatgtaa	gaagacagcc	agaaggtggc	12660
cacagccagg	aagagagctc	tcaccagaac	ccaaatctgc	tagcaccttg	ctcttgggtt	12720
ctcagcatcc	agaactgtga	gaaatgaatg	tgtgttgttt	aaaccactca	ggctacggta	12780
ttttgttgca	gcagcccaag	ctgacagaga	tagaaacaac	acaaggaccc	atcagcagac	12840
gaatggatga	tcaaaacgtg	gtgaggtcgt	gcagtgggat	attattcagc	cgtagaagga	12900
atgaaattct	gatacatgct	ataatgatga	accttgaaaa	catgttaatg	gaaataagcc	12960
aaacttaaaa	ggacaaatat	tgataaattc	cacttatatg	agttagttac	ctagaatagg	13020
caaattatgt	catagataca	gaacattaga	ggttaccagg	gttgtgggaa	gaggggtatt	13080
gtgggtacaa	attttcgggt	tggagtgttt	ttgaaaaaat	tctggaaatg	ggtagtga	13140
gtagtcaaca	tgatgaatgt	acttaatgac	actaaattgt	acacttaaaa	atggttaata	13200
ctgggctggc	gcagtggttc	atggctgtaa	atcccagaac	tttgggaggc	caagacagcc	13260
ggatcatgag	gtcaggagat	tgagaccatt	ctggctaaca	tggtgaaacc	ctgtctctac	13320
taaaaaataa	aaacaaataa	aaaaaaaatt	agccgggcat	ggtggcaggc	acctgtagtc	13380
ccagctactc	gggaggctga	ggcaggagaa	tgggtgtgacc	tgggagtcgg	agcttgacgt	13440
gagctgagat	gcgcccactg	cactccagcc	tgggcaacag	agccagattc	cgtctcaaaa	13500
aaaaaaaaaa	aaaggttgat	acctgggtgc	ggtggctcat	gcctgtaatt	tcagcacttt	13560
gggaggccaa	ggcaggcaga	tcagttgagg	tcaagagtta	aggaccagcc	tggccaacgt	13620
ggcgaaaccc	catctctatt	aaaaatacaa	aaattagtcg	agtgtggtgg	tgggtgcctg	13680
tagtcccagg	gtctgggagg	atgagcccta	ggaattgctt	gaaccagga	ggcagaggtt	13740
gcagtgtggt	gagattgcgc	cactgcactc	cagcctgggg	gacagagcga	gacttagtct	13800
caaaaaaaag	gttaaaattg	taagttttgt	tatgcatatt	ttaccataat	ctttaaaaaa	13860
tagatatata	ggagataaa	tcaacagaat	ttaataacca	gttgtaaata	gagactgagt	13920
gaggagatg	aattaaaggaa	gacattgagt	acaacttttt	ggtaggtgaa	aaactcttaa	13980
aaaaatacgt	gggcaaagat	cctacttgat	tcttataatt	taaaaatctc	ccagttagta	14040
aacaaggcta	ggtggagatt	tgcatgtgat	gtgaggtgtg	tgttctgttt	tgtaatgtga	14100
ggactgtgag	ccatctcctg	gacttgaata	tccatttagat	aattgaaaa	acggatttga	14160
gaactcagga	gacgtgcaat	gcagttaacaa	aactctgcac	ctagttgatt	tctgtctcct	14220
aatttaaatgc	ttttatggga	caaactgtta	ggcagggtgg	caagatggac	agccatattt	14280
ttgtgggttt	ctggcctgtg	ggccagcctc	agtgtcactc	ctgaggtcat	gtccaaactt	14340
agaacacatt	caggcctacc	acagtcaagg	ctccctttct	caactctagt	cctctgcaca	14400
aatatccgaa	gcctagaaat	aataatcatc	tgtccttgtg	tcttgcat	tgaaagccta	14460
ggaaagggcc	ttgggaatta	agaagaatgg	aaaaactggg	ctaactgctg	catgcttcag	14520
cttgacaggg	aatcactgaa	atggggacag	gccataaaa	gacaaccaga	agagtggctt	14580
cagcaaaagg	atcgtttttc	agagcaagct	agagaatcct	gccagcgtcc	tcaggcagg	14640
cccctgggca	cagagggttag	gcaagggagt	gtcccagcat	gttgatgcc	tgagcatcag	14700
aataatgccca	tagaggagct	tccaaagagt	tcatttcagg	ttttgtaagc	cgaacatttc	14760
taggcaaaata	aaatttgatt	ttgtgaataa	agcttgtttc	ttcaactcca	gtgcagattc	14820
tcatgatattg	atagtggctt	gtgatccaga	taaagaaaa	aatttttcaa	agattcatat	14880
tctttgtaga	tgtacggatt	tagagaccat	ctaacttaac	tcctcatc	tacagatagg	14940
aaaaatgagg	cctaaagaag	ttaagaaaat	accatggaaa	tgtcactgct	gaactgccat	15000
acgtagatgc	cgaaagaaat	tgggtaaatg	ctactgtgag	aaatacagta	ctaggtccaa	15060

```

agaatctaataacaaattaaa aatctaataatg ttattttctaa agcatccctg cacatggctg 15120
aacttacatagtttctatcttcttcttctgttgaagaa gaggcaattg gctgggtgca 15180
gtggctcatgcctgtaatcc tggcactttg agaggccgag gcgggtggat cacctgaggt 15240
caggagtttg agaccagcct ggccaacatg gtgaaacccc atctctacta aaaatacaaaa 15300
aattagctgg ctgtgggtggc cgctgcctgt aatcccagct actccagagg ctgaggcagg 15360
agaattactt gaatctggga ggtggaggtt gcagtggagcc aagatcacgc cattgcactc 15420
tagcttgat gacaagaggg aaactccatc tcaaaaaaaaa aaagaaaaaa agcaatcact 15480
aacctgtgtt gtttattaaa catgacagac tggcatgaag taattaccaa actgtaaaaca 15540
aaaaagctac aatctgccag gcatgggtggc tcatgcctgt aatccccac cttggggaggc 15600
caggttgggg gatcacctga ggcctggagt tcaagactag cctgggtcaac atggtgaaac 15660
ctcgtctcta ctaaaaaatc aaaaattagc ccggcgtggg ggcacatccc tgtaatccca 15720
gttactcagg aggctgaggc aggagaatca ctggaacctg ggcagtgggg aggttgca 15780
gagccaagat cgcaccgttg tactccagtc tgggcccaga gagtggagct cggctcaaaa 15840
aaaaagaaaa aagaaaagct acaaccttaa tctcaacttc tcataacatc atctctactt 15900
ctgattagaa gagtggaggtt attacaaaaa gactgttata ccttacacac 15960
ttctccccat gaatagtga ggtgtgagtg aaaaagacag caattttat ttttttttga 16020
aacaggttct tgcactgtca cccgggctgg agtgcactgt tgtgatcact gctcactgca 16080
gcctccacct cccaggctca agtgatcctc ctacctcagc ctctgagta gctgggacca 16140
cagttgtgca ctacctgcc cagctatttt tttttaagag atggggtctc actatattgc 16200
ttaggctagt tctcaaaactc ctggcctcaa gcagtcctcc gaccttggcc tcccaaaggg 16260
ttgtgattac aggcataagc caccacaccc agccagcagt tttagaataa agggtgagg 16320
tgctgttggg gaaatataat ttaaaaaaca aaatcttctc tcaaccaga aatcctctcc 16380
atgaaggcag tagagaaaga taagctttat tattgaataa aaattaaatg agaattgtat 16440
gcacatcaca ggcactttgc taagagatca caaagacaga aggaaatttc accattttgt 16500
acagccaagc aggtacagcc cattacatgt atgttttcta gataaatagt cctcaactaa 16560
gagaacttga cagcaccact ggtcacacag ttcattctaa ctttacctga taattgatgt 16620
gaccacttgt gttatctaag atatcaactt ttcgggggtg ggggagtggt gaaacaggag 16680
ttacttttat agcttgggtgc aaggtactca ttaagattag gctgttacc tcccacagaa 16740
actggaagat aggtatgcta tctggtaatg tttacatttc ccagatcctt gagaaagaca 16800
ttcctaggtc ataaagctga caaaaggctg attcagtttt taaatatata tatctgtata 16860
tgtatttca 16869

```

&lt;210&gt; 65

&lt;211&gt; 15000

&lt;212&gt; ADN

&lt;213&gt; Homo sapiens

&lt;400&gt; 65

```

gatctcttga tcccaggagg tcaaggctgc aatgagctaa gatcaagcca ctgcattcca 60
gcctgagtga tagtgggaga ccttgtcttt aaaacacaca cacacacaca cacacacag 120
agggcctttg accactcttg agtagaagac tcgagaagaa caaagtagaa ggccagagaa 180
gaacaaaagtt acttgaaaga tctcttatta aagagaatgt acaagctatg aaaaaaaaaa 240
aacacacaca cacacacaaa cctcatcttg aatgaaaaaa acataatgca ttgggtttct 300
ggttccttag gctgttatgg aacaacaaa gaacattatt ttgggtttctg aggtcagaac 360
tattttattc cctcaagca cactatgctt atgggttgag ggagaatgag aaataggaaa 420
ctaggaacag gctgaaatgg tctaactctg accatcta tctgcagtgt cttattctca 480
ttctaaaaga gaattgggtt attcgtgtgt cttagcataa aagtaatgat aaaaaataaa 540
gatcccgat taccagacaa taatccccta gactgtttta atgcttggtt gagtatttgc 600
ttatgatctc agactttaaa agatgggtct cccctatggg gaagcttggt aattatgtag 660
gcacatttaa tgtctgttta cttatcaaaa ttttatcatt gttagttgta ttactacttg 720
acagtccaat ttatttaatt gaaaagattg gtttaacatt tatagtcaaa gtaattgttt 780
cctgtgtttt ttctgtttta ggttattgga gtgatgagta aagaatacat accaaagggc 840
acacgttttg gaccctaata aggtgaaatc tacaccaatg acacagttcc taagaacgcc 900
aacaggaaat atttttgag ggttaagtaa ggaaatttct tcagacccat taaatgtag 960
gaaaaaatgg agctaaaaaga gctgggtggc tcacctttct catcctgtgc tgagaaatgc 1020
tggggctcac ccataagtat ccagatcccc catggacaca gggaattctg aacaaatgtg 1080
atgaaaccga tgaaatgtct ggcctgtagg tggttagtga tggagatagc ggctatatgt 1140

```

gaatcttgat	ttttgcaatt	cattagagct	ttgtaatgaa	aggaaacagt	ttgttgcttg	1260
ctttaaggat	aggttcattt	gcattttctcc	gcaaggaagt	agtaatgagt	taccaagcc	1260
tagatttcac	ccctttttga	ttctttgctg	acttaacttt	aattgaatgg	aagagttat	1320
acaaatgaat	tatctttttg	gttttttttt	ttttgagatg	gagtcctcac	ctgtcaccag	1380
gctggagtg	aatggcatga	tctcggtcca	ctgcaacctc	cgccctccag	gttcaagcaa	1440
ttgtcccgcc	tcagccctccc	gagtagctgg	gactaagggtg	cgcgccacca	tgcccagtta	1500
atthttgtat	tttttagtaga	gacgggggtc	cactatgttg	gccatgatgg	tctcgatctc	1560
tggaacctcgt	gatccgcccc	ccttggcctc	ccaaagtgtc	ggaattacag	gcaagagcca	1620
ccgcgcccag	ccaggaatga	caaatgaatt	accttataag	taaatgccat	taaggaagga	1680
tagctgggaag	atgggttgag	gggaatggag	gaccacagaa	ctagtcccta	ttaaatacat	1740
gtgcatggta	aatgatattcc	atthgacaat	aggttaatta	tctcatagca	taaggaaaat	1800
gcttaacagt	catatgcaag	atgataagct	ttcttatagc	atccaaccaa	aagatctagc	1860
cagtacaatt	tccttttgcta	tattaggggtt	agaaaggccc	ccagaggtga	accaattaga	1920
tggaatccctt	gaataaaaaca	ctggattatgc	agtgaacaga	aaaaagtcag	attgctttcc	1980
ttcttcccat	agatgtctca	gggatattta	gtttcctcag	aagataaaga	atttagtaag	2040
cgthttttttg	tgataactta	catgaaatgt	acattattttg	aattctttta	aaagaaacag	2100
ctgcatgata	acaaaaattg	tgctatgctt	gctttagctg	gtatttttgc	ctagaacgat	2160
tatctcgttc	ggacaagaag	ctatttcctaa	gaaacaatat	ttttaatcca	ggaagtthtt	2220
catthtttga	aatttatctt	actattttccc	aagcaaaaaga	gggtagttac	agattcacta	2280
agaatcatgt	gctcaccaatt	tttatttcct	aattattcct	ccttaaaaata	tattaatcac	2340
ctgacttaca	atgggtggaac	catgagtgca	tttttgctt	tattgtcaat	aacgtcttct	2400
cagaagtggag	ccacaaagggt	gcatagttct	tggaagttaa	ggctctgaatt	aagacaatcc	2460
agcataagtc	tcattaatgt	gtgattattt	tgagaaaagg	caagaagtac	ctaagaatct	2520
ccccctcact	gctccagtcc	ctgtttcatt	taagagttca	ctgtaagtaa	ctgaaaaggc	2580
ttccttgggga	ggattttattt	gaatcagttc	ttcacatgca	aaggatattg	tagaacatct	2640
cgthttttgct	ggcaggaata	tgaacatctg	ttgtgaggaa	agaaaaagt	tcatgcaaat	2700
tacactgcca	aagaagggat	gttcaagttg	agaaaccagt	gacatttctt	gtaactgtac	2760
tatgaatcacg	cgcatthttaa	tcttctagat	aatatatgga	agtgcaggaa	gggtgttagga	2820
aacgggtgttc	atthttacata	tgcgttattt	tattctgtgt	gagtgacttc	atggcaccga	2880
cattgctgtt	tttaaatgag	gatacagtaa	attgcagttc	gaggaaggct	aactggaatc	2940
aacatacccg	tagcttttaga	aagcagtttc	cgaccagcg	aagagtacaa	gagcgatgga	3000
acccctcatg	cttggaaagt	tgcacatcag	agtaaacaaa	cttgaaaacc	cctcttgata	3060
gcagaattca	cccagccttg	ttccattttc	tcttaacaaa	acacaccgca	aaagctctca	3120
caagctgctt	tgatgaagcc	acatgtattt	cccccttcac	aatttacagg	aagttactct	3180
taaaagaaaag	tgattcttgt	gtttaccgcc	tgtgttaaag	ggacagagtt	cctttttatt	3240
tctgataacag	tttgagcgaa	atacagaac	tatctgtaga	ctagcatagt	cggtacgtga	3300
gtaaggaaaa	gcaataacct	gctgtccggt	gagcacaaaa	ttcctgtctac	gaacagtgcc	3360
ttactgctgc	ttggagactg	caagtcgcag	atcacactag	gtattgactg	attgtataag	3420
gaaatttctt	aaagtctaaa	gtaaagggtg	tacctcttaa	aaagagggga	agagagaaaa	3480
ctttgtgtgg	aaggataagg	agtgtgttta	tagtttcagt	aagagtgtac	gttttaattt	3540
ttcttcttcc	tctgcctctt	tgccaagtag	cctgagtgca	tctgttatcc	agaagtagta	3600
ttactctagg	acaaacttca	aattcttcat	tctgcgttgc	ctttaaggaa	caacatactt	3660
ttcttctgtt	ctttttccaa	aaacacacgc	ctatggctct	gtgtgtgggt	ttttagccag	3720
cctcctccca	gataagggggt	tcccttccct	cctttgcatt	gaaaggaaaag	tgcaagctcg	3780
gacatgttta	tcaagaggaa	aagtgacttc	tcagtaatag	actgtcaaat	tcgggctgct	3840
gcccagagtgt	tcgctttgtt	atggcaggtg	aagttcacct	ttgccccacc	cagtgtttcc	3900
acaaaaaggc	aaggttccaa	gtattcatat	gaacaagtgt	tacttttagga	cttggagggt	3960
tggggggtgga	ggatgtttgc	atagttgaag	ccttgggctg	gggtgttagga	aacggcgagt	4020
acagaggcca	tagaaaaagc	taagactcag	tttgacgtcg	tcagccggct	tggtcttcta	4080
cccagtgact	caaagcacta	aaagtcagca	taatcggaac	tgaagtacgt	agcatcgccc	4140
atthgccatt	cactgcagta	gcaaaaagtag	tactctgtgg	tgggttaatc	ggthtgaggc	4200
agctccttaa	atgaacattt	gtgtttcatt	ttctgtttat	tttcccgaac	atgaaaagac	4260
gataaaaactg	aaatggaaaa	ggtaactgac	aaaagtgtgc	cttacctgtt	tccgccctga	4320
ttctgtctga	ttcaagacta	ttctggctaa	actgatttga	ttctttttct	aactaggcag	4380
taggggatca	gaaatcacac	acggatcccg	ctgtgtttat	tctgagaggt	gctggggagc	4440
tttggtctgt	acttcccttt	acatgcctgt	cttctctttt	ggacagatct	attccagagc	4500
ggagcttcac	cacttcattg	acggctttta	tgaagagaaa	agcaactgga	tgcgctatgt	4560
gaatccagca	cactctcccc	gggagcaaaa	cctggctgcg	tgtcagaacg	ggatgaacat	4620



```

ctacttctac accattaagc ccatccctgc caaccaggaa cttcttgtgt ggtattgtcg 4680
ggacttgcga gaaaggcttc actaccctta tcccggagag ctgacaatga tgaatctcag 4740
taagtggatt acagaacaaa aaaataaaaa atgccagtaa tgtcggttct gccctcttga 4800
actaataaca tgttgtttta ttatacggct ttgtcatgtg ttggatgaag taggtggctt 4860
aagctaggga ctaggaagag gaaaaacatt ttttgagtc cttttaacta ttaggaaact 4920
tgatcattta aaagtatata tatatatgag gagctacctt gagttttgaa ttcaggatgt 4980
tacaggaaga aatatatgtc caattctaat ttatccaaaa gcagttggga gaatttcagg 5040
gattggcca gacatgctgc gtatgcaagg tatagccctc atctgtggta ctttggcagg 5100
gcttagactg catcaaaata tttatagatg tacatttgag tgtacagtta ggatcctgatg 5160
tggaacattg taagatcatt gctagaaaaa ctttgtcata atttttcaat attattctaa 5220
gtgaataacc gtaaaagattt tacatcttag ctctcttctc tacagtaaaa aaactatctg 5280
atctcttgat cagtattata gtagccacct atcactttat ctttaacaaat tctcaattcc 5340
ttaggtttat gtgcttttac ttcttttatt tgattaaaaat tgctgtcatg acctctctct 5400
gcagagggtc gcatcatttt ggctattctc aagtgatctc tttagcaat ttaagaattg 5460
ccataagatt ctaacctctg ctgtaactat ggttgtgtgt tcttggttag accactaaat 5520
ttatttagca gttttaaaaa ttattccttt ttgttttagaa gtttaagacta aatgctgaag 5580
tttttgtaac ttttggtttt gatatcattt caaacttaag aaaacatttg aagaaaagga 5640
caaagaattt ccacttacct tttaccaggg tttaccagtt attgataagt atatccattt 5700
gctttaccag aaggctaact tgttttagtt ctcatcttca cctttgagac atttggaaat 5760
aatatcaatg ttaacataaa ttggaatttt gactttgat ttaggaccaa tgaacaagcc 5820
aagtacttac cctagtcata tataatccaa ctgtatggtt atttggattt catccacac 5880
ttcattttac ttgatctccc ttaagattgc aagattgtgt ttgcagtttt tctgaaaatc 5940
tggggctata aaagcatcag gacctcccc gtaggggagg tctgtgtgtt ggggtctcta 6000
cacaacaggt taaccttgag cttcaggaaa agaaactggc ctcagttccc cagttccagc 6060
ttaatgggtc taattagggtc ctgacaaaaa aggtggcagt tcttttccct catgtctctt 6120
cagcgctccc cgagactctg gagactctgt catatcccta gggctgagcc tcccaggaa 6180
cattcggctg ttgtggcatc tgtgtatgcc atgcccagtg ctgaggacct agtaacaaac 6240
gacaaatgca caggcacagt ggcaattttg tggaaactct attccagctg tgcgtctcag 6300
aagaagcgca cagctccctc ctggctttct taacatagtg agccacttcc acttaagggt 6360
ctccttacat tccttgagtt taatcattca ttgattcaga ggaaagtctt ttgatttttg 6420
cttttcttta aacagttcat ttgaggtgac ctaccccagt gactttgcac caaccacaa 6480
gaaaattttt tgcatgcttc ccgaccctg tgccaatcaa gggaaaggtt taaaggcctg 6540
gcgtttttat tctcaaaaga aagggttttg acagtatttt aaggttcaag tgcttctact 6600
ttgtgttcag aagcaactgt catatatact gtgaaatgac accttttatt tatccctttt 6660
tatttatgca gtatgtcccc ttttattttg gcagaatttt ttctaaatgg ttgtttaaca 6720
ttttcaagca catttcattg tccaatattc atagtaaaaga atgagagtta acaataacca 6780
gtcacattaa aacaagattc ctgctgccag ttgtgaaacc ggttgtctta ggctggcag 6840
ctgatgattg agactgtgat caggaaaatt tccactattt catcaggcct aataggtaga 6900
tttgtctccc aaatgaactg tgttgggttt ccattgctaa agcacaatag aggtggtgca 6960
agaatctcca tgagggttta aatggcagtg atggttcagg cggtagagtt tggagaagaa 7020
gggatttgaa acaaaccaaa ggaaagaaaa gtaagtagcc agaaatcaca aaatggcatt 7080
tttctaaaaa caaaggaaaa ggaataaaag aactaataag tttgaaacct ctacccctcc 7140
caaatttggc aggtgggggag gtattttttt tctatctatc taactaacct atctagaaaa 7200
cagttgacca aattatagac ttctaaatgt taactgtctt tctcagtttc agttgaaaa 7260
agactttgtt ttgcctactg cagaacttct aggttcttct ttatagtctt ggggttctta 7320
ttatagatcg aaaatgtgag tcggcataat taagccattc ggagtcttca gaagcagttc 7380
actcttgaaa tgactccgtc cgcctacagc catttaagat ttcagaacaa aaacagatct 7440
tgattttctt tttcatgtta actcaagctg ttgctgagtg ggagagtcag aaatgcacc 7500
agctccactg attactcagc tgctgaagga tgatttttta aaatgcacct ttactgtata 7560
tggaacttct aatttccacc tgtagagcat cttagggagg ctaacatgtc actctggatg 7620
ttctttttaga ataagatgca aatctatttt tctgaaggca ttagagatag caaacattta 7680
ttgtgagttt actatatact aggcactgtg ctaagtgttt tgcatagaaa gtttaaaatt 7740
ctggcttttt tgttgcccca atcataagtt tcatatcagt tcaacattca aattatatta 7800
aggtacttaa gaagaatccc tggctaaatg tgaggggcag tgccacagat ggactgaaac 7860
tttatgctta ttgcacattt atgctattat tatttgttga attatagaac caagggagtg 7920
tggaagccac tggaaaaaat atgagactta gatacataat ttgagtaaaa atggctcaaa 7980
gtcatgaggg taaagttttt tgtatttcca ttttatcga gcggcatcgt ttttaaaat 8040
cattatgaat ttgacctat atagatgttt ccaaataatt cttttccac ttcataaaat 8100

```

tccttctgt	ggctgtgaga	tgccttgect	atcagttttc	aagcttagtt	gtctttctca	8160
tcctttacca	tttttagcttt	aaaaaacaaa	agtgacaatt	agaacttcct	gcctgctggg	8220
cctcactgaa	agaccgatat	tggcctgata	aggagatat	tattttgtt	tagtggtctc	8280
agaaatccct	ctccctcagc	aagctttcca	tcacggcccc	cccgtcagca	tcttccctga	8340
tagcgttctt	ctctgtgttt	attctggggc	ttcaggctcg	cccaggagca	actgataacc	8400
gctggcagga	gataacattc	tctaaggggc	tctcaaattg	gaatcgaatc	cctcaagcca	8460
gtcagcctag	agaatacatt	taaagggttc	agttctggag	tttcacagag	ttcatttcta	8520
gacctatcag	atagcaagt	tggagtctt	tctcaactaa	attcaagcag	agacattttt	8580
tagacgatga	aggatatttg	cacaaaggct	tcagcatgat	cccccaaac	tgctgctct	8640
gaaggcatct	ccacacattg	acagccaatg	ccttcagtg	gttcctaggg	caggtgtctt	8700
ggcttgagt	actgtcctcc	aataatcaga	gtcacaacta	aacatcgtat	gttttacttt	8760
tggtttccag	gccaggctga	gcagggaatt	ttcagttttc	cctgcccaga	tgggtgtttt	8820
ttcctgaagg	catcatttat	tgtgtagcga	ggagacagg	ctggctgtgg	cagggatagt	8880
ctagaactgt	cctcattgct	gctgttccta	aatagtatct	ttaccaagta	ataacgtgcc	8940
gtctttggga	ataagtgtt	tctcttagc	ctgttctgtt	ttcttgggtg	cgctaagtaa	9000
ctcgaactcag	cttaggaagta	cctattgtgg	tttggcagag	gtgactgtca	cgcttgtga	9060
ctccaggggc	cagcactgct	gggatcctgg	ctagaccaga	cagagccttg	gtgaagtgtc	9120
taggctgtct	gcacatcgcg	aggaaggtgg	tattcacttc	gctaagctcc	ttggcatagg	9180
cagtttgaa	agggttttat	caaattcgta	ttcaacaaga	gtagaagcga	aaattgatga	9240
ctgtgtattt	cttgaaatga	gtcttaattc	ttcacattta	gttctcaggg	tatgctgatt	9300
tccttttaggt	aaaccatgaa	catcagaaag	acttttatta	acctatgaca	gggtccccc	9360
cccagtatct	ttccactcca	ttaaaatgga	agtttttttt	ttttttttct	tttttgagac	9420
agagttttgc	tcttgttgcc	cagttctggg	tgcaatggca	caatctcggc	tcaccacaac	9480
ctccactctc	cagattcaag	cgattcttct	gcctcagcct	cccaagtagc	tgggattaca	9540
gggtgtgcgc	accacgccc	gctaattttg	tatttttagt	agagatgggg	tttctccatg	9600
ttggctcagg	tggtctcgaa	cttcgcacct	caggtgatcc	gcccacctcg	gcctcccaaa	9660
gtgctgggat	tacaggcaag	agccactgca	tccagcttag	gctatcttac	tccagcctaa	9720
acagcaattt	cttatcataa	ggtctgtact	aatgaaaaca	gaatcaccca	aggctgctgt	9780
ttgtttctgt	tgtgtgcca	ttgtccgcat	tttgtgagg	aggaaacgga	actgcacttt	9840
tgagtgtgtg	gcccagagcc	ttctagaatg	agagtgcgtt	ggaagccaga	tatgtggcga	9900
ttgtgtcgcc	agctgttact	caggttttct	caagaaggag	gagcaacttt	ggcagttttg	9960
cttcagttct	ctctagccct	ctgtgtaatc	gccccctttt	ctttattttc	gcacaaacac	10020
agagcagttc	aaagcaaccg	agcactgaga	aaaatgaact	ctgcccacaa	aatgtcccaa	10080
agagagagta	cagcgtgaaa	gaaatcctaa	aattggactc	caacccctcc	aaaggaaaag	10140
acctctaccg	ttctaacatt	tcacccctca	catcagaaaa	ggacctcgat	gacttttagaa	10200
cagctggggg	cccgaagt	cccttctacc	ctcggtctgt	ttaccccatc	cgggccctct	10260
tgccagaaga	ctttttgaaa	gcttccctgg	cctacgggat	cgagagaccc	acgtacatca	10320
ctcgctcccc	cattccatcc	tccaccactc	caagcccttc	tgcaagaagc	agccccgacc	10380
aaagcctcaa	gagctccagc	cctcacagca	gcccgtggaa	tacgggtgtc	cctgtggggc	10440
ccggtctcca	agagcaccgg	gactcctacg	cttacttgaa	cgctgcctac	ggcacggaag	10500
gtttgggctc	ctaccctggc	tacgcacccc	tgccccacct	cccgccagct	ttcatccctt	10560
cgtacaacgc	tactaccctc	aagttcctct	tgccccctta	cggcatgaat	tgtaatggcc	10620
tgagcgctgt	gagcagcatg	aatggcatca	acaactttgg	cctcttcccc	aggctgtgct	10680
ctgtctacag	caatctcctc	ggtgggggca	gcccgtccca	ccccatgtct	aacccacttt	10740
ctctcccag	ctcgctgccc	tcagatggag	cccggagggt	gctccagccg	gagcatccca	10800
gggaggtgct	tgtcccggcg	ccccacagt	ccttctcctt	taccggggcc	gcccgcagca	10860
tgaaggacaa	ggcctgtagc	cccacaagcg	ggtctccccc	ggcgggaaca	gcccgcacgg	10920
cagaacatgt	ggtgcagccc	aaagctacct	cagcagcgat	ggcagccccc	agcagcgacg	10980
aagccatgaa	tctcattaaa	aacaaaagaa	acatgaccgg	ctacaagacc	cttccctacc	11040
cgctgaagaa	gcagaacggc	aagatcaagt	acgaatgcaa	cgtttgccgc	aagactttcg	11100
gccagctctc	caatctgaag	gtaggccttg	agagagagca	gtccaagggg	ctgtgagtgc	11160
atgcttgtgt	ttgtatttag	cttgcctttc	atgggtatc	gattgcattt	gcagtagatt	11220
gagcccccg	ttggggatag	tgggtatgga	ttccgcctgg	cttttgccac	ttctagtctt	11280
ttgactttgg	acaagtgact	tcccttctcc	tgattttctt	ctgaataata	aaaaaattag	11340
gggtttggac	tagaagatta	ggtgaaactc	cctgctagcc	tgtgattttt	gtgcttttaa	11400
gaaaaaaccc	attctgaaaa	catgaagatt	tcttcttttt	aagactgtct	tgatgctttt	11460
cttaagatat	ttgcatcaac	acttgagtct	tggagcagaa	atgttaggtc	tcagagccag	11520
cttgagagca	gagctaacac	atgtggcttc	ttcccaggtc	cacctgagag	tgacagtggt	11580

```

agaacggcct ttcaaatgtc agacttgcaa caagggtttt actcagctcg cccacctgca 11640
gaaacactac ctggtacaca cgggagaaaa gccacatgaa tgccagggtgc gcagtatatt 11700
ctgggtagac cttctgacct ttgtagaaaa tgcctgtgag tcacctcccc atgtcttata 11760
tagcccgtag ttaaagccaa caccagattc tgcgttgtcc catcctggac tgatggcact 11820
atggtccttc ccagtacttt gtatctgctg atgacttgag atggcacagc cagcttccag 11880
tggtgtggaa aatggtaggg gaaataaaca gccctcgtg tgctgtgtgc ccacatcccc 11940
ccgtttgctt aataccacac tggagggtgc acaaggaggc ttctcacctc ctagggtgct 12000
ggcggttggc cggtaagcct gccctccccg ttggcaactc ttaattcttc ggcttctctg 12060
tctcccttcc ctgctgtctc tctccctac actgtaggtc tgccacaaga gatttagcag 12120
caccagcaat ctcaagaccc acctgcgact ccattctgga gagaaccat accaatgcaa 12180
ggtgtgccct gccaaagtca cccagtttgt gcacctgaaa ctgcacaagc gctgcacac 12240
ccgggagcgg ccccaacagt gctcccagtg ccacaagaac tacatccatc tctgtagcct 12300
caagggtcac ctgaaaggga actgcgctgc ggccccggcg cctgggctgc ccttggaaga 12360
tctgacccga atcaatgaag aaatcgagaa gtttgacatc agtgacaatg ctgaccggct 12420
cgaggacgtg gaggatgaca tcagtgtgat cctgtagtg gagaaggaaa tctggccgt 12480
ggtcagaaaa cgaaaagaa aaactggcct gaaagtgtct ttgcaaagaa acatggggaa 12540
tggaactctc tctcagggt gcagccttta tgagtcata gatctacccc tcatgaagt 12600
gcctcccagc aaccactac ctctggtacc tgtaaaggtc aaacaagaaa cagttgaacc 12660
aatggatctt taagattttc agaaaacact tattttgttt cttaaagtta gacttggtag 12720
gtcaggggtc ctgtaggaag tggctgttac ataaccag ccttgcaaag ctctctcgac 12780
agcaaatggt tccccctcac ctctggaatt aaagaaggaa ctccaaagt actgaaatct 12840
cagggcatga acaaggcaaa ggccatatat atatatatat atatatctgt atacatatra 12900
tatatactta ttacacctg tgtctatata ttgccccgt tgtattttga atatttgtgt 12960
ggacatgttt gcatagcctt cccattacta agactattac ctagtataa ttatttttc 13020
aatgataatc cttcataatt tattatacaa ttatcatc agaaagcaat aattaaaaaa 13080
gtttacaatg actgaaaaga ttcttgttaa ttgagtata aatgtatttt tgtcttgtg 13140
ccattctttg tagataattt ctgcacatct gtataagtac ctaagattta gttaaaacaa 13200
tatatgactt agtcacacct ctctctctaa taatggtttg aaaatgaggt ttgggtaat 13260
gccaatgttg gacagttgat gtgttcattc ctgggacct atcatttgaa cagcattgta 13320
cataacttgg gggtatgtgt gcaggattac ccaagaataa cttaaagtag agaaaacaga 13380
aagggaatct tgtatatatt tgttgatagt tcatgtttt cccccagcca caattttacc 13440
ggaaagggtg caggaaggct ttaccaacct gctctccct ccaaaagagc agaatcctcc 13500
caccgccctg cctccccac cgagtcctgt ggccattcag agcggccaca tgacttttgc 13560
atccattgta ttatcagaaa atgtgaagaa gaaaaaaatg ccattgttta aaaccactgc 13620
gaaaatttcc ccaaagcata ggtggctttg tgtgtgtgcg atttgggggc ttgagctctg 13680
gtggtgtttt ttgtgtgttt tttttttt tttttttt atgtcaaaat 13740
tgcacaaaca tgggtgtcta ccaggaagga ttcgaggtag ataggctcag gccacacttt 13800
aaaaacaaac acacaaacaa caaaaaacgg gtattctagt catcttgggg taaaagcggg 13860
taatgaacat tccatcccc aacacatcaa ttgtatttt tctgtaaaac tcagattttc 13920
ctcagtattt gtgtttttac attttatggg taatttaag gaagatgaaa gggcattgca 13980
aagtgtttca acaacagtta cctcattgag tgtgtccagt agtgcaggaa atgatgtctt 14040
atctaattgat ttgcttctct agaggagaaa ccgagtaaat gtgctccagc aagatagact 14100
ttgtgttatt ctatctttta ttctgctaag cccaaagatt acatgttggg gttcaaagt 14160
tagcaaaaaa tgatgtatat ttataaatct atttatacca ctatatcata tgtatatata 14220
tttataacca cttaaattgt gagccaagcc atgtaaaaga tctactttt ctaagggcaa 14280
aaaaaaaaaa aaaaaaaaaa gaacactcct ttctgagact ttgtttaata cttgtgacc 14340
tcacaatcac gtcggtatga ttgggcaccc ttgcctactg taagagaccc taaaaccttg 14400
gtgcagtggt ggggaccaca aaacaaccag ggaggaagag atacatcatt ttttagatt 14460
aaggaccatc taagacagct ctattttttt ttgcccactt tatgattatg tggtcacacc 14520
caagtcacag aaataaaaaa ctgactttac cgtcgcaatt tttctgttt cctccttact 14580
aaatactgat acattactcc aatctatttt ataattatat ttgacatttt gttcacatca 14640
actaatgttc acctgtagaa gagaacaaat ttogaataat ccagggaaac ccaagagcct 14700
tactggtctt ctgtaacttc caagactgac agctttttat gtatcagtg ttgataaaca 14760
cagtccctaa ctgaaggtaa accaaagcat caggttgaca ttagaccaa tacttttgat 14820
tcccaactac tcgtttgttc tttttctcct tttgtgcttt cccatagtga gaatttttat 14880
aaagacttct tgcctctctc accatccatc cttctcttt ctgcctctta catgtgaatg 14940
ttgagcccac aatcaacagt ggtttttatt tttcctctac tcaaagttaa aactgaccaa 15000

```

<210> 66  
<211> 46340  
<212> ADN  
<213> Homo sapiens

<400> 66  
tattttactt cagtaacaga aatgaaaaga aatgttttaa tgttgctgat tgtattacct 60  
tcaggatcaa tagcagaagg acaaacttct ttgaggagat ctcctagtgt gtgcaactgt 120  
ccatctgcag ccacaggacg aaacagcttc tgaatgaaag gtcttccagt cgttgtctat 180  
ttgaaaaagg aaaaaatgat tcaagcaatt aagtccttgt tgctgccaat tacaaattta 240  
tatatcataa actttatgtt ggcathtagt gccctttgat acgggtgttag cataattaca 300  
caacatcaca gatgtggtat cactgtgaaa aatgtttaac atgataaatt caggtaaatc 360  
taattctgag gaaacagaca aatccaaagt tgggtgggac attctaaga taattggctg 420  
ggacccttca aaaacttaaa gacattaaaa agcaaacacac acaaaaagat atcaacaaaa 480  
gcatTTTTtC tcagtatctc tttaaagagac taacaaagca aatacaaaac ataaaccatg 540  
gctgaatact aaattgaaga aggacatttt ttagaaatcc aactatgaaa cacagttttg 600  
ggataaatgg ggaataacag aatggacaac tgataatatt attgagttaa tgtcaaattt 660  
cttaggtaca ataaggacaa tccttatttt taagaaattc attgttcaag tgtttaggaa 720  
agaagtgccca tgatatccaa aacttaatct tcttctctt tttttggaga cagagtctcg 780  
ctctgccacc ccggtggag tgcagtggcg cgatctcagc tcaactgcaac ctctacttct 840  
caggttcaag tgattctcat ggctcagcct cccaagtagc tgggactaca ggagtgcgcc 900  
accattgcca gctaactttt tgtattttta ctagagatgg ggtttacca tgttgccag 960  
gctggtctca aactcctgag ctcaaggcaat ctgccggtt cggcctccca gagtgttagg 1020  
gttacaggcg tgagccaacc gctcctggcc ccaaaactta accatctaatt ggttgagaga 1080  
gagacagaga gagagagaaa gagagagaca gagaatgtgt gtgtgtgtga agacaaagca 1140  
aaaaataaaa aatattaact aatggtgatt ctaggtagag ggtgtatgat tttagtagt 1200  
tcattatttc aacttttcca taggtttcac aatttccaaa acagcagatc cagccatttc 1260  
atctgacaaa aactgttagc agcactacat cgtaatttat tgctaataat ctcatgttt 1320  
tactcttaaa attgtttcat ttactaaatt tccttagtga tgatggaggc tttatcatga 1380  
cagagtacag aggctctgaa atgagccagt gtctatgaag agcaccactg ttgcaagtc 1440  
catgatctt taccagtt tcctttatct gttaatttgg gacattccat atctcttgag 1500  
tttgtgtgg aaataaatga gcaactttgc caaccacaga gtaataaaat aaatgttaaa 1560  
gagaataaaa gcatttttac ctctctctc cctcttaacg gttatttcac ttaagatgg 1620  
taaattttaa gctttctgag atgaaaaact attaaaactt aacaagaaca gagaaatgcc 1680  
atacatcat attttttgt tgctgttct ctgagacaag gtttcaactc gtcaccagg 1740  
ttgaattgca gtggtgcaac cccaagttg caatctcca cctaagctc cagagttagt 1800  
gggactacag gtgtgagcca ccatgctcag ctaattttt tactttttg tagaagggg 1860  
tctcactatg ttgcccaggc tgcctcatat tttaagaa tatgacttca aacacttagg 1920  
cattagcgac aaggttttgt ttttgtctt taatgacaga ggtatccct aacatatttg 1980  
acacaactgt tagagatttg gtttaaaaag aaatagacat ggatgaagct ggaaactatc 2040  
attctcagca aactaacaca ggaacagaaa accaaacacc tcatgttct actcacaact 2100  
gggagctgaa caacgagaac acatggacac aggcagggga acatcacaca ccaaggcctg 2160  
tcggggagta gggggctagg ggagggatag cattaggaga aatacctaac gtagatgagg 2220  
ggctgatggg tgcagcaaac cccatggca catgcatatc tatgtaacaa acctgcacat 2280  
tctgcacatg tattccagaa cttaaagtat aatacaaaat gaaaaataa ataaaaataa 2340  
gtagaaaaaa taaacatgta agcatgtgag ctgcctttcc taattctatg ttatgtatt 2400  
cactgaatac atagtatttt aaaatagtaa tccaataata tatttgatg tttgtgaca 2460  
gtatgaaaat tgtaattttt aaaaaatctt gataatatgc attgaatatg atttaattca 2520  
cttactatt tgaactctt agggattatt tttaaaaata tgattgatat ctttgatat 2580  
gttttggctc tgtgtttcca tccaaatctc atctcaaatt gtaatccca cccgtctagg 2640  
gaggactgt aatccccatg tgcgagggga gggaggtgat tgggtcatag ggggtggtt 2700  
cctcatgttg ttctcgtgat actgagtga ttctcatgag atctgatgg tttaaaagt 2760  
gcagtttttc ctgcaactctc atctctctt cctgctggct tgtgaagggt cctgcttccc 2820  
tttctgccat gattttaagt ttcttgagg ccccaagaac cataggaac tgtgagtc 2880  
ttaaaccttt tgctttata aattatccag tctcagatat ttctttaag cagagtga 2940  
acagactaat acattcttca atttaaaaag ccatactttc tcatacaagt tgaaacca 3000  
aacaatatca tgcataatca agtgattaac tgtgtaaga taataaggtt gaggagttca 3060

gagaagaaaa	gaaatgaata	gggaactgta	gtgataaatt	aaaatagcca	tccctcactc	3120
aggggtttttg	atcttcaggc	catgaagaag	cttttaaatgc	tttttagcaa	aggaagtaat	3180
gttggtgaaa	ggctttttct	gacgactaat	ggaaagcagt	gctatgtaag	gtgacttggg	3240
tatgaaccaa	aaccagaatg	actggtgaga	ggctgactga	atacagcaag	cttatgtgaa	3300
gacaactgga	gctggtgcag	tggaaaagga	agacagcagg	actgtaccca	caactcaaag	3360
aaaaaagtca	gaaggtacct	cccgcagtcc	aacctgaaaa	caacaaagtc	aaaggaatct	3420
tttcaagaat	ttggagctct	cattcatatc	ctaattagtg	tatgaaatgt	gaggtgggct	3480
tgtataatg	aaattacctg	gaatatttct	aacacaaaga	aataataaat	gcttgagggtg	3540
gtgaatatcc	tcatttgatc	attacacatt	gcatgcttat	agcaaaaagt	tacatgtacc	3600
ccataaataa	ttgcaactat	tatgtatcca	taataattaa	aactaaaaag	ttaaaaatta	3660
cctgaaaaaa	aatgctaaac	aggaaaggcc	aactagtctt	ggttacataa	taaaaaacag	3720
aaattcttct	ctaacctcac	tattggagaa	atatcctggt	atttttataa	atcttttttt	3780
tcaccctttc	ccaaatctga	gcaagtatta	taaagggtata	accttcaaca	atctttttatg	3840
atgaggtatt	tgcttactgg	ggacaaagcc	ccagtgtctat	tacatagtgt	agctaaacgc	3900
tgtagaatgg	taaaaacaag	aaaatgctca	gcaaagtgtt	gtttctcatt	taatgaaaat	3960
cttatttttaa	aacacaaaaa	ctcaatatac	cccaaccaaa	aatctgatga	acattttctg	4020
tttaatatatt	attatacagt	acctttaaaa	acgtaatat	cttattctta	aaaatttagt	4080
gtgctagcaa	atagcaatta	agtacctaag	tcaatcagga	cgacaaaaaa	atactcaatt	4140
tggggagtta	gttacttcta	tcattctgaat	gcgtccctcc	aaaattcatg	ctgaaaccta	4200
ttcctcatca	tggcagtatt	aagaggtgaa	gcctttgaga	ggtaattagg	tcagtggggc	4260
agagtcctca	agaaatggat	caatgctctt	ataaaagagg	ccccaggagg	cttgtaaggc	4320
ttttgccctt	tctgccatgt	tgggggggtg	gggggtgggg	cgagcaacc	agtgtcaact	4380
ctgaagcaga	gagcagccct	caccagaaac	cgaatctgtt	gaagccttga	tctctgactt	4440
cccagcctcc	agaactgtga	gaaataattt	tctgtgtgtt	ataaattacc	cagtctaggc	4500
tgggcgtggt	ggatcacctg	aggctcaggag	ttcaagacca	gcctggccaa	tatggtgaaa	4560
ccccatctct	actaaaaata	cagaaaatta	gctgggcata	gttgtgggag	cctgtaatcc	4620
cagctactca	ggaggtctgag	gcaggagaat	cacttgaacc	cagaaggcag	aggttgagct	4680
gaatacaagt	catgccattg	aactccagcc	tgggcaacaa	gagggaaaact	gtctcaaaaa	4740
aaaaaaaaaa	aagtacacac	tctaaccatat	tttggtatag	cagcccaaat	ggaatggact	4800
aagacaatta	cccttaaaat	aaaagctccc	atagagagat	catgcattca	agtacagagg	4860
ttcttaaggg	caatgggaat	ggaggacata	ttcctgcaaa	cttttcaaca	gctctcatta	4920
gcccgatgtt	agagctctgc	aaagaagact	aaattatact	gagaaatatt	tttaaatctc	4980
cacaaatagg	aatgctgtaa	acgttgattt	agtatatata	aaattagaca	agactaacaa	5040
tatccaatgc	aatctaaatc	ttaggttgac	agacaagaaa	gccactgcaa	acaggaatat	5100
accacaatac	ctgatcttgc	cacatatttg	taaatatgca	aagtatttca	ataacttcca	5160
agaaacagta	tactctcat	gagaaataac	atgatgtaag	tcaccttga	aactgtcctt	5220
gttacttttt	caaagtatg	ttagtcatat	cttaacacca	aatgaaatga	aaaactgagg	5280
tggtaatggc	tggctgctcc	catctctcct	ctactcatgt	gccttcacca	atacagcaat	5340
cattttttct	tatatgggaa	atttacagtg	ttgatatagc	tcagagatat	attgaagaaa	5400
agcagaaaaa	cgaaacttat	aaacatttta	ggaaacctta	tgtattttct	taaatagttc	5460
aagtgtaaaa	cttagaattc	ttataaataa	tgtgtgttac	agctatatgt	taaatgggtg	5520
ctcatgcctg	taatcccagc	acttcaggag	accgaggtgg	gaggagagct	tgagcccagc	5580
agtttgagac	tcacccgggc	aacacagaga	gacctcatct	cttaaaaaaa	aaagaaagaa	5640
agaaagaaat	gaaatgcaaa	gaaaaagtct	ctatttcaaa	tgtagccagt	agagccaata	5700
ggttaaccaa	tattaacatt	aacgttgata	aaacaagaaa	tgatgattta	ctataagctg	5760
aaaatcagac	aatgtatgga	ctttaagagt	aacaggcacg	atcatcacia	acttaaatca	5820
ggtttgagtc	ctatgagtta	tatacagtta	catgatgcaa	caaaagatgc	cagccagttg	5880
ttaaagagta	ttagattcgg	ctgggggtgg	tggctcatgc	ctgtaattcc	agcacttttg	5940
gagggcagg	agggaggatc	acgaggtcgg	gagtcagaga	ccagcctggc	caatatagtg	6000
aaacctgac	tctactaaaa	atacaaaaac	tagtcaggca	tgggtggcacg	tgctgtgaat	6060
cccagctact	cgggaggctg	aggcaggaga	attgcttgaa	cccagggggc	ggaggttgca	6120
gtgagccgaa	atcgcgccac	tgcactctag	cctgggcaac	agagcaagac	tctgtctcaa	6180
aaaagagtat	tgaattcaag	tcctgtttct	gtcatttatt	atggaaacat	ggacacaact	6240
acctatcttt	cctgaacctc	agttttttca	actgcaaaac	aggaatatat	acatatgtgt	6300
atatatacat	ctgtgtaaac	acatatgtgt	atatatacat	ctgtgtaaac	acatatgtat	6360
atgtataaat	ggagataata	cctacattat	agttttctgag	ataataaaa	gcacaacaca	6420
attctgacac	ataacaattt	gtaacttaaa	acataccatc	accagggcca	ctagttttat	6480
aacactgtaa	tgcatagtct	aatttaatac	tatgcaaaact	gtgttcactc	aaggttttat	6540

tctcttttaa	tttcattcat	ttactcttca	gttgtttgta	agctaaaaag	tccagaatca	6600
tgaaatctag	aagtttacgt	tttaattgtt	tcttatatgg	caaggaaaaa	aaaaaggcca	6660
aagtcatttt	aacactactt	tcaaaatcag	cctagaactt	aacactaaaag	gcatgaccca	6720
taaaaaggaa	tactaataaa	tagacttaat	taaaattaaa	caacaacaac	aacagctaa	6780
cttttgttct	gcaaaaagac	ctgtgaagag	aatgaaaaaca	taagccgcag	gctgggagaa	6840
aatatttgca	aaccatattt	ccgagaaaag	tcttgtgtct	ataatatata	agaactcccc	6900
aaattccaac	gttttttaaaa	aaagcaata	atccaattag	aaaatgggca	aaagacatga	6960
acagacattt	taccaaagag	aatatatagg	tggcaaataa	gcatatgaaa	acatatctca	7020
cacatcatta	gccatttaag	aaatgcaaat	taaaaccaca	atgtgatatc	attacacacc	7080
taccaaata	tccaaaataa	aaattagtgg	taacaccaa	tgctggtgcg	catgtggaaa	7140
aatagtcctt	cacacactga	tggatcaaat	gcaaaacagt	acagtccctc	aggaaaggag	7200
tatggcagtt	tcttacaaaa	ctaaacatgc	acttaccata	tgaccaagta	attatactct	7260
tgaatattcc	cagaagtaaa	aatgtcttct	ccaaaaaact	tatacatgaa	cgttcatagc	7320
tgttttattc	gtgagagtca	aaaacagaaa	gcaatcccag	ggctacccat	taaaaacagg	7380
gaatgtctat	aaactgactg	taatatgtct	gtcccacgga	atactactca	gcaataaaaa	7440
ggaacaaact	actggtatat	gcaacaactt	ggatagatct	caagggagtt	atgttatgtg	7500
aaaaaagtca	atctcaaaa	gttacacact	gcatgactcc	actgatataa	cattagttaa	7560
atgacaaaaa	ttttagaaat	ggaaaacaaa	ttagtagttg	tcagagggtta	gggaagaaa	7620
gcagtaaggt	aggttgctgt	gggtatcaaa	gggtagccta	agagatcctt	ctgttgaaac	7680
gggtatatat	tgaatatagg	gtgaatttac	atatgtgata	aagattgcat	agaactaaa	7740
acacacacac	agtatatgta	aaactaagga	aatctgagta	aggtttgtgg	attataattaa	7800
tacaatttcc	tggttgtgat	actgtactgt	aattatgcaa	gatgttagaa	tgggggaaa	7860
ctagatgaag	gggtatgtag	tctttctgta	ttatttctta	caattgcatg	tgaattctga	7920
attatctcaa	aataaaaaat	tttttcaaaa	tttcaaaaaca	actagtctag	agctttgtta	7980
atcaaagttt	tctctgagga	cctgtagcat	tttggttatc	acctggatct	tattaaaatg	8040
tagattctca	ggctgcatat	tggaaatcct	gaattggaat	ccgcatttta	acaagatttc	8100
caagtgtatt	attgttaaa	tttgagaagc	actagtctac	aacaatgact	tttaacgttt	8160
caacctactc	taacacactt	gaaggccata	acaaaattca	catcaataac	agttgctcgg	8220
ttggacagtg	actctcaaca	caaataagtg	aggaaagggtg	gggactcaag	actcaggtag	8280
caggaaaagc	cccttaggtg	atcctgatga	aatgttttct	ccatcctggc	tgaaaaacc	8340
agaacagta	attaaaggctc	aaaacaaaag	taatgtttat	aatactggag	atcttataaa	8400
ggcagataat	atatactata	acagagcaaa	ggtaattatt	acaatgtata	aatcttataa	8460
gaaccaaata	cagaattaaa	atcactaagc	acataatgaa	aatcctttta	aaagtataaa	8520
aatgaatgta	gtctaagtaa	atactaataa	tggcagttat	agtgaagaaa	gctctagagt	8580
cttttacttc	tatacttcc	catctatttc	caaaactgac	ccttcgttat	8640	
tcaataaatt	tatggcctgg	tacagtaata	agagcatgat	atttaaagcc	agtcagaaga	8700
cacatattct	agctctggat	ggcacttgat	gacgatggat	tcagcttatg	gttccaatcc	8760
cagctctgtc	aattagtacc	tatatgacct	tagtcaata	cttaaacctt	cttgtgttac	8820
ttgtgtgtca	attgtatcat	ctataaaatg	aggatattaa	cagtatatata	ctcatagatt	8880
tttttgtgaa	ggttatacaa	ttaattcata	taaagtattt	agaacaatgt	ctagcacagt	8940
gaattctcaa	tgagtgttat	aattgttctt	tttaaattgtg	acttgactct	caacagaact	9000
ctactgaatt	ctaatatgta	ttctgtattg	agctgtcaaa	aaaaataagg	attataataa	9060
catatactat	tcttgtagtc	aaccctgtta	ctatgttatt	actagtgtca	gttttgttat	9120
tttggtcata	catattgttt	tacatacatt	aagaattatt	agaaatgttg	gtttattaaa	9180
aatgaccatt	tatggctaga	agggtatata	tctggctcac	tgactgtgga	gtcaatgtcc	9240
ataaagagga	ggaagaatgc	catcagagta	aaaggagatt	ctattcactg	aaacaaagt	9300
ataaaaagct	atgaaagaga	aaaacataaa	ataaccaaag	gggtgaaact	taacagattc	9360
ccagtagatg	cacaatgcac	tgggtgttaa	aacttaaaat	ggccttaatt	aaaagccaa	9420
cacggatgga	ggtgctgggg	gagtcctcta	cggacacagc	aggcagaatg	taacaatgac	9480
aaagggctca	agtttatatta	aaaagagatt	ggacaggccg	ggcgtgggtg	ctcacgcctg	9540
taatcccagc	actttgggag	gctgagggcg	gtggatcatg	aggtcgggag	ttcaggcca	9600
gcctggccaa	catggcgaaa	cctcatctct	actaaaaata	aaaaaaatta	gccgggagtg	9660
gtggcggtga	tctgtagtcc	cagctactca	ggaggctgag	gcaggagaat	cacttgaacc	9720
tgggaggcaa	aggttgcaat	gagctgagat	catgtcactg	cactccagcc	tgggcaacag	9780
agtggagctg	ctcaggatct	cccaaagacc	caaactccctg	taactgaat	gcataatatc	9840
atattatatt	gtgaggttta	gtatgttgta	tttttcaagg	tatagcaaca	agtttttatt	9900
catcagctac	tttgtgtgtg	tgctttgttt	ttaagtcttt	tgaacagga	tgggtattta	10020

ctacatttat	aagtaaaatt	tatttgattt	acaagggttg	cttaagtgtg	tcacaggatt	10080
tcacttgtaa	tatttgcagg	tgcttaaaaa	atcagctata	ctaaactata	actggaatta	10140
gcaaagtcca	tttattgatt	aatcaagaat	ataattagat	ttgcctaact	atataagtag	10200
tactatgtgt	tatttaagaa	ttaaatctag	aaaagggatg	gactctggaa	atatcaagaa	10260
gtgaaaaaga	ctgctctcat	ttttgtacaa	caattactaa	atcttctaagt	agcattaatt	10320
gaactgaaaa	ggcatttttag	aaaaactaga	ttttacaaat	tataactcta	ataaaacaca	10380
actaaactatg	agtgtgcttg	ttcatgccca	aaagctacct	tccaaaatta	aaaaccttat	10440
tggtatggctg	ggtgcagagg	ctcatgcctg	taattccagc	actttgggag	gccaaaggcg	10500
gcggatcacc	tgaggtcagg	agttcgagat	cagcctggcc	aatatgggtg	acccgtctct	10560
aacaaaaata	caaaaaattag	ccggggcgctg	tggcggggtgc	ttgtaatccc	agctactcgg	10620
gaggctgagg	caggagaatc	acttgatcct	gtgaggcgga	ggttgagtg	agctgacacc	10680
gtcccactgc	actccagcct	ggcgagagc	ccagagcgag	actccgtata	ttaaacaaaa	10740
caaaacaaaa	ctcaaaaaaac	cctattggca	attactaggg	ccatcaaata	agtataattt	10800
cacttgacac	acaattttga	gataatgaac	cgaacttact	atttttgaaa	atattacata	10860
ataaatattat	tggaagcttc	attgctgaaa	tggtgacaaa	gatgaatagc	aataaaaactt	10920
ttcttataga	tcttttagcaa	aaacaaaaaaa	accccaagca	tactatggta	cattacttta	10980
gagaatcaag	tagctgctag	ttgagtaata	gtggtaatag	gcactacaat	gatataaaca	11040
aattacaaca	aagaatattg	tttttatttc	ctgtccatgt	tttaaaaaag	ctttgggtttt	11100
gaggtctggt	caacaaaagca	taggtacaac	aacgactact	actactaaca	tataagtagc	11160
ctggatagaa	ttatcttaat	agtagtacct	aaagtgcagga	tctctaagta	atgatcagaa	11220
ggcaggaata	aatttttatca	gaaatcttca	ttcattacat	atttactatg	catttaccag	11280
ggtatcacta	tgctaattgga	tacaaagata	aataacatgc	aaacaactgt	aatacagtgt	11340
tatgtgataa	cagaaatag	tacaaagcac	tatgaaaaaaa	attacaaagc	ttgagcaca	11400
attttaactc	tggaacttact	ggcattttaga	gcaaaaaccaa	aacaatccta	actgggtta	11460
ttcattttct	aagagttgga	agctaratca	gtaggtacaa	agtaaaatat	gctaattgtg	11520
gtagaaagta	aaatattaca	acagtagaga	atttcaaaaag	aagataaaaa	taatggagg	11580
aatatagaag	gtcttcaagc	ttccagcttg	aaatcacatat	ttttttttaa	atagagaaag	11640
agataaagtc	atttgagtat	tcagaggcca	gactgaatat	aatggtaact	ctgagaaatc	11700
agtggataag	gagagaaaag	tggaactaaag	gccatagcat	atagagcttg	gaatgtcaaa	11760
tgtagtgga	ataacaaagg	tttgggttga	atcccaactc	ccaacaacgt	actgtgtatc	11820
tagagcaaat	tacatcaacc	ttggggagta	ctgtttctga	atctgaaaaa	tgaggaaaaa	11880
ttatctttga	acaattgatg	tgataattaa	atgagatata	tgaaatatct	aatgtaacaa	11940
gtgcttaaca	atgactagtt	cttttcattc	ctctcttgaa	ccattgtgaa	acgtagaacc	12000
aagaaaggta	acagtattta	gttgttacag	aaccatttaa	gagagaataa	aaaataactg	12060
gtattctaac	ttcagtttcc	ttgaagtct	tgtaattgag	aataaatatt	atgtggcaca	12120
aagaaaaaga	aaacaggggg	ttacacagga	tatgctgcca	gactttacca	acaatgacac	12180
atgatattctg	cttcaactgt	cccattgcata	tttggcttaa	gatataattca	tgcatatcaa	12240
attttcatatc	acatggtttt	caaaagaaga	ttcattaaaa	ttagcttaag	aatgtacaca	12300
atatacaata	cctcattaaa	taaaaagaac	agaccatttc	caaatgaatg	cttttagagc	12360
tttacagtaa	acagtctttt	ggtggtagaa	agagggggaa	cagagagggg	agtgggtggg	12420
agtctgtagc	acttatcaga	ctacttttat	cctttatgta	gagaaatagg	agagttgaaa	12480
ataagcactt	tctgtactta	tgttgagagt	ctgaagccca	cttttaatat	tcttgacaac	12540
actaaaaaat	aataattaac	atttgaaaag	ctgtcattat	tatagtcagg	gacactta	12600
ctccaaagga	gaagtttctt	aattgatact	atgattaaat	aaaagcatcc	atcagaatta	12660
tatccacaat	ctgggttgga	gtttatgttt	tgtcttattt	aaattgttat	acttattata	12720
attctgtcta	gacagtccca	aatgtacttt	gtcatacaaa	cacttgaggc	aaattttctt	12780
caaataagcg	caacactttg	tttcctcttc	gtatcctttg	actgaataac	gtgtggtaca	12840
gagaagtaat	acttcccttt	cttgggatcg	agatcaatct	gatgcttggt	ataagcccat	12900
ttacagaaca	aatggtattg	cttttaaaatt	tttatatgaa	cttatcagta	gactagccaa	12960
aaaagaagct	tcatataaaa	gtgctaggat	tgatattctt	agtaataatt	aggtaaaatt	13020
tctaaaaatt	tctcccaaaa	gatctgaaaa	atcataccaa	gggaagtata	gtttaaaatt	13080
cattatatat	aatagcttta	aaatatcttt	gctaattcta	cccaagccca	cactaaaaag	13140
actaatacaa	aaagaatgta	attaataaac	tattttcttc	tgaagaatca	aagggcactt	13200
ctgcatatga	acatgtttta	tccttttggt	gtacttacat	aaaataatta	agaaacactt	13260
tttaattagta	taaacaaaga	aatcaaaaata	gcaagaagaa	atgtctgagt	aaaagcagct	13320
gtgctgacct	caaaaagtga	attctgttct	cttgatgcc	agttaaagtgt	ctaaccagg	13380
gaaaagtgat	tctaaacctg	ggctaggagc	tagtgagagc	cttcaaacag	tctcacctac	13440
cctcaccct	caaggaaatg	tctatgggtt	ctgtgggtgaa	cgctaaagtt	tataacatgg	13500

```

gaatatttat tattttgttt ctaacacaaa taatttttaa aaattttatt tactaaagt 13560
acatcaaagg gaaatttcac aaaaattcct ttgaaatttt tagaagtag aaataaagg 13620
aagtgataaa tattttacag atttcaccac ttacgtaatc tgatcaacaa attttaaaa 13680
catagcactt gaatactatt aaaaatataa taaaaaggta acatagtaaa actataaaa 13740
tctttaaaaa aaatataaga ggaaaccttc gtgaccttgg attaggaat ggttcttta 13800
atacggaac ctaaaaatac aagcaaccaa agaaaaaac agacaaactg gacttcatca 13860
aagttaaaaa cttttgttct tcaaagaca tcatcaagaa aataaatccc acagaatgg 13920
acaaaatatt tgcaaaccat atctgataag agaccactat tcagaataag taaagaatt 13980
gtaaaactta taaataaaaa gttaaagaag tcaattttta aatgagcaaa ggatctgaa 14040
acaattctcc taagaaatac gaatggctag ttaaattgcat gaaaagatg ttgacatca 14100
tggtcattag gaaagagcaa aaaccaaaat gatatactcc ttcataacca ctaagactg 14160
tgtaattaaa actatagaaa ataagcgttg gcaaggatgt ggacaaattg gaacctctcc 14220
catacactga tggtagaaat gtaaaatggt gcagatgctt tggaaaacag tctgacaata 14280
cccaaagggt ttaaacgttg aattaccatg caaccagca attctactcc taagtatct 14340
cccaagagaa atgaaaatat atgttcacca aaacatttgt acataaata taactgcag 14400
ttttattcat aatagccaaa aagtggagac aatccacatg tctatcaat ggtgaattg 14460
taaaaaaat gtgtatctt catacaacta ttactgggcc aaaaaagaa tgatgtatt 14520
atacatgcta caaaatgaat gaaccttaaa aacaatatgc aagcaaaaag aaccagaca 14580
aaaaggccat atattacatg atgttaatta cataaaatgt ccagaaggga gaaataaat 14640
agtagttgcc aagggtctga gggaggggga atgatataag tgactgccaa tgggcatgg 14700
gtttcttttt aggtgatga aaatgttctg aaattttatc acgggaatgg ttgcacaac 14760
ctgtgtaact tagaattcag tgactcctaa aaccaatgaa tagcatgct taaaagggt 14820
cctttgtcga gcatagtggt tatagtctta gctacttggg aagctgaggc aagaggata 14880
cttgagccag gagtccaggt ctgtactgca ctatgatcat acctgtaaa agccaccata 14940
cacaccagcc tgggcaacac agaccatgtc tctaaataaa taaacaaata aataaataa 15000
agggtagcct ctgtagtatt gagattatac ttcaagtaag ctgttattaa aaaaaaaaa 15060
gttatcatat gggtagcagg ggaatcatt ctgggatgat ggctaactc atcagtatt 15120
gatttatacc tatgcatcat accttatggt tgttttatgc attttgtggg ttttttaaa 15180
aaattatatt tcataaaaac aaattttaaa aaaattaaag tcaagaaccc caaaacaaca 15240
aagatcagag atacatttct accttatcaa ttcagaaaaa ttacaagtt ttttcttaa 15300
aattgtatag catcatgggt attttaagtt acctgtagga atttaataa ctttgtctt 15360
actgttcacc aaaaactcatt taatttcatt gttctgatac tgaaaaatgaa gctgaaaagt 15420
tttgaaaatta caatatgcta gtttaaaaag gtttactaaa atacataatt tcattataag 15480
gagtaatatg aaataaaagt atcaaatatg ggaccattaa aaatgtctc actaacaat 15540
tgctacccac attgtggact cactgcgtcc actgtttgag agcttttcca gaacgctcg 15600
caccagttag gttagccaag aactcctcat cttcactttc ttcctcacra gcttggaacc 15660
tctggattcc caccacact gctgtgacct gaatggggaa gagaaacgc atagtaagg 15720
aactcttctt tttatagatt tctgaattag aatctggcat tacaaaaaga caatgttata 15780
aatccaggtc agagtttata gttctatttc actattactt atatggcttg tctaggaa 15840
ttaactatta tttacaatgt aagtacctat ttccacaaaa aaattcaaaa ttttggaata 15900
caatatctga agagagaatg gtctattgaa tccaaagtag gctgatacat cccaacagta 15960
tttcagattg agataataat aataccacca attcatcaag tcaaatata tgcttattt 16020
ccacaatgga agttttaaaa tagtataaac attttaatat atagcaggct taacttatga 16080
ttattaaaca gggttctaag aaaatagtat acatcaaata ttaatgtgct tcttgataa 16140
tttaggtgac aatttatcca tctgagaaat gcaaaaagaga ctttggttag ggttgagta 16200
aggagcattc tgtgtcaaaag aattcactag caaaagaggg tatactgtag ttacaagcta 16260
taatcactgt acttatttta aatccctctt cagaaccagg tcttaaaaag tgataaaca 16320
ggcctctaga ataactatca accaaactat agaaaagagt gcaagagtgt ggtgttcta 16380
cttaaaatat ggtgttttat tcaataaatt ttatttaagg ctccaaaagc agcagcctca 16440
ttccccagaa atcatagtta aatgaaatct tcttactaaa aggaaaaatg aatcacata 16500
tttaacgtga acattttaaa aacactctaa agcaacaaa ctattcaat gtatgtgata 16560
tggcttagaa agcatgtag gtaaaaagga ctaaaaactc taataatggt tgggccaaaa 16620
gtaaatttgt tagttctact ccattaagca ttcctcaagc agtgtaaaaa tcagagtta 16680
agttacactt tgatgtgtag atcctttgaa agccactcta ccctgtttta tatgaagca 16740
ccgcagctaa aatgaacacc tagtgaagag tatgaatgct gcaatacata agcagagctc 16800
agaattgtcc caagctgatt ctaagttaact ttaaacatgt atgcagagt agaatatgac 16860
ttacttctta gaagtaacag ataattacct ttggcataat gaaaaaaact ttaaatgtaa 16920
gttaatacag gtattttccc tttagcaaaag ctttgctttt aaaagaaaac ttcaaaact 16980

```



aaattaaaat	aggaaatgct	ctactatgta	gtaaaaatac	tttttagatt	actgaagcaa	17040
agaaaaggaa	ggattctatg	agggaggaaa	agtgggagaa	aaatgtaaag	aaaaaaagga	17100
agaaggaaag	aaaagagaaa	aggaggaaa	aacacaagga	cagaaaggcc	tattgaaata	17160
tattatttct	ttcaaatttt	aaacgagcag	aataaattct	tttgttttat	aactatgaaa	17220
taatctatgt	tcctcttatc	tatgcttgga	aaatttagac	aaaatgttaa	gagtaagtac	17280
tacatttgat	ttccgggtct	tcagctctga	aaacaagctg	tttcttaaca	tacgtcaatt	17340
ttctatatatt	catgtcattt	ctatttgcaa	atgttataaa	gttcaatatg	atgtaaaaca	17400
tggttaaatg	aagttcaaaa	ataagtataa	catacattag	tttggctatt	ccaaatttca	17460
tgcacattaa	ctcagccaca	catctaacac	agtcagccct	ccctatccag	gggttctgca	17520
tctgcagatt	caactaacca	tgggtcgaaa	atgtttttgt	accaaactatg	tacaggcttt	17580
ttttcttggt	atcattccct	aactacagta	taacaactat	tttcacagt	tgtacatgtg	17640
tattgaaatat	tataagtaat	ctacagataa	tttaaagtat	acaagagggt	atgcataggt	17700
tatatgcaaa	tactacacca	ttttatatca	gactctcaaa	catcagtaga	atttggtaac	17760
ccaggagggt	cctggaacta	atcacccaga	ggatctcgaca	gatggctata	tataaatcac	17820
tcagtgaatt	caggattcac	attatttcac	aactagtata	attttatggt	gttcacataa	17880
ttgtgtcaca	cagatatacat	gcagacaggt	gactttcatg	aaaagattac	acccaagata	17940
gacatatggt	ctactcaaat	acgggtttcca	aatgtgtatc	caatcttggt	taattataat	18000
caaactcacc	attccattga	taagcgacct	ctaccaacct	gcttatcccc	tccaagcaat	18060
ataacagtgg	ttctctgaac	caatattgac	cctcctttaa	attgatagcc	tttttttaaa	18120
aagctaacca	ttgagaagta	catactgttg	aagacagaac	atattctgta	aaatgctccc	18180
aagatatcaa	agtcagatga	tacaactgaa	tgtttatgct	agattatatt	tctaagctga	18240
gaattacatt	ttaatatacc	ataagcaatc	tgcataaagaa	gcaacttgcc	taaagatttc	18300
aggagtcca	agtatgcata	tgtcaatatc	tgtatcaata	tgtaatatca	atataatcaa	18360
tgcacacaca	aatacgtaac	tgtacttata	tcactctcct	agcactaatt	attacaacaa	18420
atctgcatgc	actgcaaagc	aaaagtataa	tataaaatcc	caaaaaacct	tgaaaattta	18480
ataaaaccaa	aaaacaggca	tcacacacaa	gaactgaggc	gtatacttca	ttaatgagta	18540
tgtatcctgt	atatgaaatg	tcaaacaaaa	ttaccaggc	tcagggttaga	aataaagata	18600
ggacattagt	ctttgtattt	ttaaattgat	ttttcttctt	aattattcct	aatgataacc	18660
ctatatatta	cctacttaaa	attatttagca	aatagttatt	ttaaaagtat	gagtaattag	18720
accaaagca	actctcatat	ttacccaaaa	gaaggaaacca	ctaccaagaa	tcaaagccta	18780
gtaattctgt	tcttaacaga	cagggtgtgt	gtattctggc	atgttacatg	aaaatcactt	18840
atgagaagaa	cagaaaaaaa	aattagaagg	tagttttcac	tatggaaata	ggtaagtgtat	18900
taagcagatt	ttcttacacc	atgaaattgt	cagcagactc	aataatcacc	ctaaggggca	18960
tcattctgga	tgccgacatt	ctctatgatg	gaaagggact	gaaagtataa	tgcactaatg	19020
acataaagaa	accaatatcc	aatagtaaa	ttgaagaaat	aaacattctt	tggacaggaa	19080
gtacagtgaa	ctttgcaact	accaagaatg	tattatgcca	gcagtaaat	aggaaactaa	19140
agcccatgtc	aaccaatgaa	aaatgggagg	actgaaatca	atcattaaag	cagcagcaag	19200
gttctaacta	ttctaaggta	taggctacct	ctggcgtata	ttatcagagt	tgacaattct	19260
tccaagaaat	cttaacatca	actgtaatt	gaggctcttt	aaaaaataat	ataaaccagg	19320
cagtagactt	acattttgta	atattttctt	ctaagagctg	tacattaaga	ttttatttgt	19380
gatataaata	ctatcaaata	attagctata	gaacagctct	attttcaaca	gttataacat	19440
tttaagccat	ctcacattta	acctaaactt	ttatcaaagt	tcaaaaactga	ggccgggtac	19500
ggtggctaac	acctgtagtc	ccagcacttt	gggaggccaa	gatggggcga	tcacttgagc	19560
ccagggaattc	gagaccaacc	tgggcaacat	ggtgaaaacc	catctctata	aaaaatacaa	19620
aaattagctg	cgctggtgg	tgtgcgcctg	tagtcccagc	tactagagag	gctgagggag	19680
gagaatcacc	agggcctggg	agatcaaagc	tgcatgtagc	tgagatcgtg	ccactgcact	19740
ccaccctggg	tgacagagtg	agacctgtc	tcaaaaaaaa	aaaaaaaaag	aaagaaagaa	19800
aaaaaaatca	aaactgatca	cttgaggtcc	aacttatgtt	tactatatct	acttatattc	19860
ccaaagacat	cttaaggaga	gatgaaatca	taaaaagggt	aggatgagaa	agaaaaatgt	19920
aagttagtaa	ggtcaatttt	tacatatatt	aggctagcat	aataaaaaata	tgagtgtctt	19980
attattattt	ttttttgaga	cagagtcttg	ctctgttgcc	caggctggag	tgcatgtggt	20040
caatcatggc	tactgcaat	gtctgccttc	caggttcaag	caatccttgt	gcctcagcct	20100
cctgagtagc	tgggattaca	ggtgtgcgtc	accctgccca	gctaattttt	gtattttcag	20160
tagagacagg	gtttcaccac	gttaaacat	gagtttgccc	aggatgggt	caaactccca	20220
aagtgttagg	attacatgcg	tgagccactg	cgtctggcct	aaagtgtctt	attataacca	20280
agaattttatt	tgtggagaga	ggtaaaagaaa	actcattttt	agtgaataaa	ttaaaactgc	20340
atcattcaca	atctatcttt	caaaatgagg	tattaactat	tttggcttct	aaaattaccc	20400
catatactac	atgcatgagc	atgggaattg	aagttatttt	attcctaagt	ttgagacttc	20460

tagcaataac	agatgggaca	aacaagctaa	ccaaaaaatt	aaaagaaaaa	cctgggaaa-	24000
aagaaatcca	aagggggtct	gaaaaagtct	aacatatttc	tgataatcca	gaaagccata	24060
cacatgtata	gagctgtgta	cacgctcaaa	aaacatctac	gaaggcccta	aactctcacc	24120
tatgggaaac	cctgaggctc	tgtacaagaa	gaaagtaaaa	tccagttata	aattgcttgc	24180
cgtatcattg	aaggcaatgc	cccaacattc	acacataggg	ccctggcaaa	gattggaaga	24240
tactctagtt	ctaggcattc	aagaaaaatc	cttctaatac	tcagatgata	actaaactca	24300
ccaagcagta	acttttaggg	cctgtgtgat	aaaaaataaa	aacctgaaa	aattagttca	24360
ggaaagaaa	taaaccaagc	acagcaacaa	caaaaaacaga	ccttgggaaa	ggggggaagc	24420
atctgggttc	cagagttatt	ctgtttatac	atataaaaata	ttcaggtctc	aacaacaaca	24480
aaattacaaa	gacatgcaaa	gaaacaagta	taagccacaa	actgggggga	aaaagcagca	24540
gaaactggcc	ctgaaaaaga	ccagatgctg	gactttactg	acaaagactt	taagagagtt	24600
atttttaata	tgcgaaaaga	actaaaaaaa	agtttatcta	aagaactaca	ggaaaagtac	24660
agaacaatat	ttctgatact	tcagaagaac	cactttttgt	cactacagat	tagttctgtc	24720
tgggtctagaa	cttcttaaaa	acagaatcat	agagtataat	ctctttatac	cagctcttcc	24780
tactcaacac	aatgttgtgt	gagatttata	catgttgttg	catgtatcat	tcccaaacag	24840
aaatagaaa	tatagagata	aataggagtt	acaaaaaagt	accaaaacaa	aattctggag	24900
ttgaaaagca	caaaaactga	attaacttga	ggggctcaac	agctgatttg	ggcagccaga	24960
agaatgaatc	agcaaatcta	aagataggtc	aattgcgaga	aagagaggga	agaagggaag	25020
aaggaaggaa	aggaggctca	gagacccaag	agacaccatc	aggcatacca	atatacata	25080
aatgagagtc	ccagaagaag	atgcagaaaa	agggtcagag	tatctgaaaa	aataatggcc	25140
ctaaacttcc	cgaacttgac	ccaaaaaatt	aatctacaca	tccaagaaga	taaacaaact	25200
aaaaagaata	aaatcaaaagc	gatccacacc	taggtacatc	ataatcaaat	gactgaaata	25260
taaagagaga	ctctcaaaaac	aggcaaggga	cttatgtata	aaacatcttc	agattaataa	25320
caaattttct	atcagaaatg	atgttgtcaa	taggcaatca	gatgacataa	tcaaaagcact	25380
gaaagaagta	gaatgtcttg	gacctggaat	gctgggtggc	acctgtaatc	tcagtatttt	25440
gggtggccaa	ggtgggagga	tcacttgagg	caaggagtgt	aagaccagcc	tgggcagcag	25500
aaagaggctc	tgtctctaca	aagaataaaa	agattggctg	aatgtgggtg	tgtggacctg	25560
tagtcccagc	tgtcaggcgt	gctaagggtg	aaagatcgct	tgagcccagg	agttggaggc	25620
tgcagtgagc	tatgactgtg	ccactgcact	cttgcaagtg	agaccctgtc	tctataaaga	25680
aaaaatgtca	acaaaaaact	acatgcagaa	aaactgcact	tcaagaaatg	atcagtacct	25740
tgaagctctg	aaggtgctta	agactgtaga	tcaataccat	agaaaataat	ttagtattta	25800
ggaatgtaga	aaaatttaaga	cagccttggt	tgataactac	acataatact	gtaactgttc	25860
ttgcactggt	ctggttattg	tcaagctatg	agcacaaaact	gatgactgaa	atacagaata	25920
cagaacagga	tataaaatct	tatcaggtaa	agttaggcaa	gcaattacta	gttgtaattc	25980
aacttgaagg	agaaggaata	aggaaccaac	tcaaaccagg	cagcaatgaa	ttgtaaaaaa	26040
gcttaaggta	aaacaaacag	ggaaataaaa	caactcagaa	cctaagcata	tcgttaagaac	26100
ctaacttaac	aaggaggggc	ttaaactgat	tattttacag	cttgggtgca	attatcccac	26160
aaaaaacttt	caggagtttc	accagtccat	aaactatttg	gttattagaa	aatagcttta	26220
ttgggtacc	ctctttgggt	cccctccctt	tgtatgggag	ctctgttttc	actctattaa	26280
atcttgcaac	tgactcttc	tggctcgtgt	ttgttacggc	tcgagctgag	ctttcactct	26340
ccatccacca	ctgctgtttg	ccgccatcgc	aggcctgcc	ctgacttcca	tccctctgga	26400
tctagcaggg	tgtccgttgt	gctcctgatc	cagtgaagag	cccatgtccg	atcccgaact	26460
ggctaaagac	ttgccattgt	tcctacgcgg	ctaagtgcgc	gggttcaccc	taattgagct	26520
gaacactagt	cactgggttc	cacggttctc	ttctgtgacc	cgtggcttct	aatagagcta	26580
taacactcac	cgcgtggccc	aagattccat	ttattggaat	ccatgaggcc	aagaacccca	26640
ggtcagagaa	cacgaggctt	gccatcatct	tagaagcagc	ccgccaccat	cttcggagtt	26700
ctgggagcaa	ggaccccctg	gtaacaattt	ggcgaccaca	aagggacctg	aacccgcaac	26760
catgaaggga	tctccaaaagc	ggtaatatgt	gaccactttt	gcttgctact	ctggcctatc	26820
ccttagaatt	ggaggaaaat	actgggcacc	tgtcggcccg	ttaaaaacga	ttagcatggc	26880
cgccagactt	tagactcagg	tatgaggcta	tctggggaag	ggctttctaa	caaccctcaa	26940
cccttctggg	ttgggaacct	tggcttgcct	ggagccagct	tccactttca	attttctctg	27000
ggaagccaag	ggctgactag	aggcagaaa	ctgtcgtccc	gaactcccgc	cattagccgg	27060
ttgagatcat	gtcgcagcca	gaagtctcta	ctcaacagtc	gccccatgct	gcgctcctac	27120
cttcccttct	gtcccacacc	tcctgggtcc	caaccacgac	tttcttgaaa	gtgtagcccc	27180
aaaatttctc	ttacctctga	atctacttcc	tctgatcccc	gcctcctagg	tactaatggt	27240
tgagactttc	atttccctcta	gcaagtgtga	tctccaaaag	gatctaagga	agctctatgc	27300
tgcgccctta	ggcatctagg	ctataaaccc	agggagtctt	gtccctgggt	tccctcctga	27360
tttaggtata	cagctctaga	catgggcagt	tatgtgggac	ctgttcccc	ccacccttgc	27420

cagggcccca	agtttgtaaa	tggctaagag	aggaaacaga	gagagacaga	gagaaagaga	27480
cagtgaagaga	cagacagaga	cagagagaga	gagagacaga	gaggagagag	agagagacag	27540
ggagagacagg	gagagagaca	gagaggagag	ggagagagac	aaagaggaga	aaagaggcaga	27600
gagacaaaaca	gggagtcaga	gaaagaaaaga	caaagataga	aatagtaaaa	aaaaacagt	27660
tgccctattc	ctttaaaagc	cagggtaaat	gtaaaacctt	taattgataa	ttgaaggtct	27720
tctccgcgac	cctataacac	tccaatacta	ccttgttgtc	agcgtaaaca	agggcgtagc	27780
ctgaaaacac	taagaccact	gacaacccat	agccttccta	tcaaaaatcc	ttaacatcca	27840
gtgacctgcg	gatggcccaa	atgcattcaa	tctgtagcgg	caactgcttt	gctaaccaga	27900
aaaagtagaa	aagtaacttt	tagaggaaac	ctcattgtga	gcacacctca	ccggttcaga	27960
attatttctaa	gtcaaaaaag	caaaaaggta	gcttattaac	tcaaaaatat	taaagtatgg	28020
ggctattctg	tcagaaaaag	gtaatttaac	actaaccact	gataattccc	ttaacctgac	28080
agatttcctt	acaggggatt	tcaatcttaa	ttaccataca	aagggtccgac	cagacctagg	28140
aggaactccc	ttcaggacag	gatgatagat	ggttcctccc	aaatgactga	ggaaaaaac	28200
acaatgggta	ttcagtaatt	gatagggaga	ctcttgtgga	agcagagtta	gaaaaattgc	28260
ctaataattg	gtctcctcaa	atgtcagagc	tgtttgcaat	cagccaagcc	ttaacgtact	28320
taccgaatca	aaaagactat	ctcaatcctg	actcaaaagc	ttacttatac	cctctcga	28380
acgaatttgc	ctaagaactg	ttgtttatgg	gaatgcattc	tgatggagca	gctgggtgt	28440
tatgaataac	tcaggaaactc	agcctagctc	taggactcac	ccctgagcac	aaaggcaatg	28500
ttgggcacgc	tggtaaaagga	ccactagaat	ccagcagccc	ggaccctttt	ctttgtgatc	28560
aagaaaggcg	ggaaaagggg	tgagggtcgc	tacatcagtg	agcataacta	atccgataag	28620
cagagggtcca	tggttggtta	cacaccccg	aaaggaataa	gcattaggac	catagaggac	28680
gctctaggac	taatgctcat	cggaaaatga	ctaagtgtgc	tggtacccct	atgttctttt	28740
ttcagatagg	aaacgttccc	ctcaaggcaa	aaacacccct	aagatgtatt	ctggagaatt	28800
gggaccaatt	tgactctcag	atgctaaaga	aaaaaagaca	tattcttctg	cagtaccgac	28860
tggaacgat	atactcttta	agggggagaa	acctggcatc	ctgagggag	cataaattat	28920
aacaccatct	tacagctaga	cctcttttgt	agaaaagaag	gcaaatgggt	tgaagtgtca	28980
tacgtacaaa	ctttcttttc	attaagagac	aactcgcaat	tatgtaaaaa	gtgtgattta	29040
tgccctacag	agaccctca	gagtcctacc	ccctacccca	gcatccccc	gactcctcc	29100
ccaaataata	aggacccccc	ttcaacccaa	acggtccaaa	aggagataga	caaaggggta	29160
aacaactaac	caaagaatgc	caatattccc	cgattatgcc	ccctccaagc	gggtgggagga	29220
gaattcggcc	cagccagagt	gcacgtacct	ttttctctct	cagacttta	attaaaaatag	29280
aactaggtaa	attctcagat	aaccctaatg	gctatattga	tgttttacaa	ggtttaggac	29340
aatcctttga	tctgatattg	agagatatata	tggtactgct	aaatcagaca	ctaaccctca	29400
atgacagaag	tgctgccgta	actgcagcct	gagagtttgg	cgatctctgg	tatctcagtc	29460
aggtcaatga	taggtcgaca	acagagggaa	gagaacgatt	ccccacaggc	cagcaggcag	29520
ttccctagtg	agaccctcac	tgggacacag	aatcagaaca	tgagagattg	tgccgcagac	29580
atttgctaac	ttgcgtgcta	gaaggactaa	ggaaaactag	aaagaagcct	gtgagttatt	29640
caatgatgtc	cactataaca	cagggaaaag	aagaaaatcc	taccgccttt	ctggagtgtg	29700
taacggaggc	attgaggaag	catacctctc	tctgtcaact	gactctactg	aaggccaact	29760
aatcttaaa	gataagtta	tcactcagtc	agctacagac	attaggaaaa	aacttcaaaa	29820
gtctgcctta	ggcccggaac	aaaacttaga	aaccctattg	aacttgga	cctcagtttt	29880
ttataataga	gatcaggatg	agcaggcaga	atgggacaaa	tggtgataaa	aaaaggccac	29940
cgcttttagt	atggccctca	ggcaagcgga	ctttggaggc	actggaaaag	ggaaaagcta	30000
ggcaaatcaa	atgcctaata	gggtttgctt	ccagtgcggt	ctacaaggac	actttaaaaa	30060
agattgtcca	aatagaaata	agccgcccc	tcgtccatgc	acctcgtgtc	aagggaatca	30120
ctgtaaggcc	cactgcccc	ggggacgtag	gtcctctgag	tcagaagcca	ctaaccagat	30180
gatccagcag	caggactgag	agtgcgccgg	gcaagcacca	gccccatgca	tcaccctcac	30240
agagccctgt	gtatgcttga	ccattgacgg	ccaggaggct	aactgtctcc	tggaactgtg	30300
tggtggccttc	tcagtcttat	tttctgtcc	cagacaacgg	tcctccagag	ctgtcactat	30360
ccaaggggtc	ctaggacagc	cagtcactag	atacttctcc	cagccactaa	gttgtgactg	30420
gggaacttca	ctcttttcac	atgcttttct	aattatgcct	gaaagcccaa	ctcccttgtt	30480
agggagagac	attctagcaa	aagcagggcg	cattatacac	ctgaacatag	gagaacaccc	30540
gtttgtgtc	ccctgcttga	ggaagggaatt	aatcttgaag	actgggcaac	agaaggacaa	30600
tatggacgag	caaagaatgc	ccgtcctgtt	caagttaaac	taaaggattc	tgccctcttt	30660
ccccacaaa	ggcagtagcc	ccttagaccc	gaggtcaac	aaggactcca	aaagattaa	30720
gacctaaaag	cccaggcct	agtaaaaagca	tgcaatagcc	cctacaataa	tccaacttta	30780
ggagtacaga	aacccagtgg	acagtggagg	ttagtgaag	atctcaggat	tatcaatgag	30840
gtcactgtcc	ctctatacct	agctgtacct	aacccttata	ttctgctttc	ccaaatacca	30900

gaggaagcag	agtgggttac	agacctggac	cttaaggatg	cctttttctg	catccctgta	30960
catcctgact	ctcaattctt	atttgccttt	gaagatcctt	caaaccsaat	gtctcaactc	31020
acctggactg	tttcacccca	agggttcagg	gtagccccc	atctatttgg	ccaggcatta	31080
gccaagact	tgagccggtt	ctcatacctg	ggcactcctg	tcctttggta	tgtggatgat	31140
ttttactttt	agccgccagt	tcagaaacct	tgtgccatca	agtcacccaa	gtgctcttaa	31200
atcttctcgc	tacctgtggc	tacaaggttt	ccaaacaaaa	ggctcagctc	tgctcacagc	31260
agggttaaata	cttagggcta	aaattatcca	aaggcaccag	ggccctcagt	gcctattctg	31320
gcttatcctc	atcccaaaac	cctaaaagca	ctaagaggat	tccttgacat	aacaggtttc	31380
tgccaaatat	ggattcccag	gtacggcgaa	atagccagac	cattatatac	actaattaag	31440
gaaactcaga	aagccaatac	ccatttagta	agatggacac	ctgaagcaga	agcggctttc	31500
caggccctaa	agaaggccct	aacccaagcc	ccagtgttta	gcttgccaac	ggggcaagac	31560
ttttctttac	atgtcacaga	aaaaaacaga	aatagctcta	ggagtcccta	cacaggtcga	31620
tgagcttgca	acccatggca	tacctgagta	aggaaattga	tgtagtggca	aagggttggc	31680
ctcattgttt	atgggtagtg	gcggcagtag	cagtccttagt	atctgaagca	gttaaaataa	31740
tacaaggaag	agatctgtgt	agacatctca	taacgtgaac	ggcatactca	ctgctaaagg	31800
agacttgttg	ctgtcagaca	accgtgagga	aagtaactaa	aatcgtaaat	ccccatggcc	31860
ctcccttatc	atatttttct	ctttactgtt	ctcttaccct	ctttcactct	cactgcaccc	31920
cctccatgct	gctgtacaac	cagcagctcc	ccttaccaa	agtttctatg	aagaatgcgg	31980
cttcccagaa	atatgtatgc	cccatcaaat	aggagtttac	ctaaaggaaa	ctccaccttc	32040
actgcccaca	cccatagcc	ccacaactgc	tataactctg	ccactccttg	catgcagtca	32100
aatactcatt	attggacagg	gaaaatgatt	aatcctagtt	gtcctggaag	acttggagcc	32160
actgtctgtc	ggacttactt	cacccatact	ggtatgtctg	aggggggttg	agttcaagat	32220
caggcaagag	aaaaacatgt	aaagggaagta	acctcccac	tgacccgggt	acatagcacc	32280
cctagcccct	acaaaggact	agatctctta	aaactacatg	aaaccctcca	taccatact	32340
tgccctggtaa	gcctatttaa	taccaccctc	actgggctcc	atgaggtctc	ggcccaaac	32400
cctactaact	gttggtatgt	cctccccctg	tatttcaggc	catgcatttc	aatccctgta	32460
cctgaacaat	ggaacaacta	cagcacagaa	ataaacacca	cttccgtttt	agtaggacct	32520
cttgtttcca	acttggaat	aacccatacc	tcaaacctca	cctgtgtaaa	atttagcga	32580
actgtagaca	caaccaactc	ccaatgcate	agggtgggtaa	ctcctccac	acgaatagtc	32640
tgccctacct	caggaatatt	ttttgtctgt	ggtaccttag	cctatcggtg	tttgaatggc	32700
tcttcagaat	ctatgtgctt	cctctcatte	ttagtgcctc	catgaccatt	tacactgaac	32760
aagatttata	caattatggt	gtacctaagc	ccacaacaa	aagagtactc	attcttccct	32820
ttgttatcgg	agcaggagtg	ctaggtggac	taggttctgg	cattggcggt	accacaacct	32880
ctactcagtt	ctactacaaa	ctatctcaag	aactcaatgg	tgacatggaa	tgggttgccg	32940
actccctggt	caccttgcaa	gatcaactta	acttccctagc	atcagtagtc	cttcaaaatt	33000
gaagagcttt	agacttgcta	acctctgaaa	gagggggaag	ctgtttattt	ttaggggaag	33060
aatgttggtta	ttatgttatt	ttagcggaa	aatgttggtta	ttatgttaat	caatcctgaa	33120
ttgtcacaga	gaaagttgaa	gaaattcgag	attgaatata	acgtagaaca	gaggagcttc	33180
aaaaacacca	gaccctgggg	cctcctcagc	caatggatgc	cctggattct	ccccttctta	33240
ggatctctag	cagctctaatt	attgatactc	ctctttggac	cctgtatctt	taacctcctt	33300
gttaagtttg	tctcttccag	aatcaaagtt	gtaaaagctac	aaatcgttct	tcaaatggaa	33360
cccagatga	agtccatgac	taagatctac	cgtggacccc	tggaccggcc	tactagccca	33420
tgctccaatt	gtaatgatat	cgaacgcacc	cctcccagg	aaatctcaac	tgcaaacacc	33480
ctactatgcc	ccaattccgc	aggaagcagt	tagactggtc	gtcagccaac	ctcccaca	33540
gcacttgggt	tttctgttg	agtgggggga	ctgagagaca	ggattagctg	gatttccctag	33600
gccgactaag	aatcccaaag	cctagctggg	aaggtgacca	catccacctt	taaacactgg	33660
gcttgcaact	tagctcacac	ccgaccaatc	aggtagtaaa	gagagctcac	taaaaatgcta	33720
attagacaaa	aacaggaggt	aaaaaaatag	ccaatcatct	atcgctgag	agcacagcgg	33780
gaaggacaat	gatcgggata	taaacccagg	cattcaagcc	ggcaacggct	accttctttg	33840
ggtccctcc	ctttgtatgg	gagctctctc	tgtcttcaact	ctattaaata	ttgcaactgc	33900
aaaaaaaaaa	tagcttaatt	gaagaataaa	ttaatacaat	aaaaggaaata	cattttaagt	33960
atacagttca	aactgtaaca	gtgttacagt	ttcaagagga	ccccttcaac	aagatatgtg	34020
gcatttccat	catgccctaa	aagtttcttc	ttgtccctta	ctggttgggt	ccatctctac	34080
tacacctcc	tgacctggcc	cagaccttgg	cctcagaaga	atcatttttt	tgctactaca	34140
tattagtttt	gtctgttcta	gaacttctta	aaaacagaat	catagagtat	gttctctttg	34200
tattggtttc	ttttactcaa	tgtaatgttc	tgtgacattt	atccatatta	ttgcatgtat	34260
tattcctttt	aatcctgaat	agtatgctgt	tttaggaata	taatgcaatt	gtttattcat	34320
ttacctgttg	acagatatct	gagctattat	gatggatatt	atgaataatt	ctgctatgaa	34380

cacttctgta	caatgttttc	tcggacatat	attttcattt	ttcttgagtg	gagctgttag	34440
aactgttggg	tcagaaaagta	agcatatggt	gaattttgaa	agaaaactggt	aaactcttct	34500
ctaaagtgat	ttgtaccatt	ttacactcct	actaataatg	tatgagagtt	atatttgcct	34560
cacagccttt	ttactacttt	gttaatcttt	ttagtactgt	caaccttttt	aatttatctc	34620
atctagggaa	cgtgaagtag	tatctcactg	ttattttcat	tttcttgatg	agtaacaata	34680
tcgtgtatct	tttcatgtgc	ttattagcca	ttcctatatc	ttttgtgaaa	tagttaactc	34740
aaatttgtaa	ctaaagggtg	tttcttgagt	ttcaggtagt	aagcctattt	ccctcaagtg	34800
aataaactac	agtcttgtaa	tgaaaaaatta	aacacagtg	agacattttt	tgtataagtt	34860
gttttactct	gtgtatgtct	ggtttgctta	gtctattatt	atatgcccc	tgaagcaaaa	34920
cacagtgtct	atttctacta	tgagtatcac	tagcacatag	aactgtgtct	gccccaaagc	34980
tgaactcaat	aaatatgtta	atgtgtatgc	atgcacatac	atctacatgc	atgtacatct	35040
atacacacat	ataaacatat	atataatttt	agacccacaa	atctaaagaa	actaattctc	35100
gagcctctgg	tttgaagaat	tctcaaaatta	ttaacatata	tttatgttcc	actccacata	35160
cactgtacct	gaaatagccc	tactgttcta	ctttggtaaa	tcaggcaaat	ttaatttttt	35220
aaataattaa	gattccaact	aattttaaaa	tataatttga	aagttaacaa	tgaatatcat	35280
taacataaaa	agtaatttta	aataaaaagc	aaactaaacc	caataagagg	aaagaaaggt	35340
gggtgtatt	tctttaatcc	tttaaaaatt	aaatcacaca	atgctccaat	gaaatcttca	35400
tttaactga	caaactatgc	ccatgaaaag	tctcatatgc	aactgtctaa	acctcaataa	35460
acataattcat	cttcttgcaa	aaaagatatt	tctttataat	atgcacatgc	agtatatata	35520
attttgaggc	agatttgtac	tttagtcctt	gttccattgc	ttaccggctg	gctgtcctct	35580
gtctgtcat	tgacctccaa	cttaaaaaat	aatacttgcc	ttgtctaccc	cacagaagtg	35640
ttatgaaagt	caaacaaggt	agcataaaag	tattttacaa	gatataaagt	gctataatac	35700
agattttaaa	aatcactcta	catcccataa	tactttgttg	tacaatttta	gagcaatagt	35760
agaaaagagc	aattttggc	taattgaaaa	tccagtcctg	aattccataa	aatgtatgat	35820
atgaacatta	tagtacatca	tattacgagc	cccaataaat	cactgcttat	atagttgggt	35880
aggatttcct	tagtttggtc	atatagttta	tataatttat	cagtccttat	tttgtgagag	35940
gcattgtgag	gagcataaag	acataagcac	agtagagagc	cttagcttct	ctacatttat	36000
taaaagaagc	aatctcttgg	gtattttaat	aataatttaa	gtattctggg	aagaaatgaa	36060
attaacttca	tagactgacc	ttagattact	atcattacaa	aaagatgcct	gagtgtatct	36120
tctttaacat	accagtattt	atctttataa	tgttatattt	acttgaatca	gaagtgaagt	36180
ccttttaagc	actaagcatc	cattctatac	tttcttgtct	ttacatatga	gatacaaatc	36240
atatttttaa	aacttttatt	tacttttatt	ttttagagac	ggagtcttgc	tctgtagccc	36300
aggctggagt	acagtggcat	gatcttggct	caccacaatc	tccacctcca	cttcccaggg	36360
ttcaagtga	caaatacatc	ttttaagcac	agattctcaa	catgtatcct	agcatgctac	36420
tgccataact	agggtgtgaa	ttaagtatta	aagacagctt	accccaataa	ttactgtaac	36480
atatactctc	aatgaaaaaa	gaacatatta	acactatac	ttggatggga	ttctggggag	36540
taaccatctc	ctctctcccc	tttcttccaa	attccatctc	ctattaacac	accagctctc	36600
ctgagctaag	cagctcctgg	ggttggggaa	gggtgtacat	ggagaaaagc	agaacctcta	36660
cagtgttttc	ctctctggga	ggaactagca	ggcatagcaa	cagaaaaagc	tgaataaaaag	36720
ctgtaactct	ttctattcct	gaggcagaca	gagagaagac	cagggaacaa	agagacttct	36780
accaagagcc	ctgccaggta	ttgatacctt	tgatactgag	aaaatatctg	ggatatgaaa	36840
tacaaatgct	aaataagtat	ctttgaaata	ggggtaaaaa	aataaagggt	cttgatgagt	36900
aaaatgggta	gtatttttta	ataacctgat	aatgagcttt	aggaaaaggg	aagggtcaacg	36960
ttatggaatg	aaaacacaga	ggtaccaaat	ttaaaagcat	aaaaaaaagt	ggaggggggg	37020
aaccacaata	cttcatcaaa	ctagcaataa	acttagtatc	atttctaatt	agaaaacgct	37080
gaaggaaatc	acttagatct	gataaagact	aggctataat	tctaactgat	gaaacactta	37140
aactgtatca	attaatacca	gaaaacaaac	acagaaaagt	ctactagaac	catcattatt	37200
cagcacagtc	ttggtaatgc	aatactataa	tagcaatgca	ataaagcaag	aaaaaaaaaa	37260
gtttgtaaaa	acacaatagg	atgagatttt	tgtttttcca	atgccataaa	taactagaaa	37320
tggaacaaaa	ataaagaaaa	acaaaatcta	caaaacacct	ggaaataaaa	agaaaaatgg	37380
tctatttgaa	gaaaacctta	aaatctatgc	agaacataaa	acaaaatctg	aataaaaaag	37440
aatatcatgt	tcttgtctgg	gaagacttaa	tatcataaga	aagtgaatta	tatcaaaaat	37500
taaatcgaaa	tttaattgat	ttccatctct	aatcagacag	gacactatgg	ggaactgaat	37560
aagtgtattt	aaaagtcatg	gaaaattaat	aactgagaat	aacctgaaa	agtatgaaaa	37620
aaggagacaa	atgaattgct	ccaacagata	tcagaacgct	aaaattaaat	aaaaatacta	37680
ctaggataag	aaaatacata	tactgatgta	atgaataaag	aatccagaat	tagattccag	37740
taagtcaaac	tactttacta	taaacacagg	gtggcatatt	catccagtg	gaaaaggaca	37800
gtaagaagtg	agtaaaactat	ggcccaactg	ccaaattgtg	gcctctgcct	atttttgcaa	37860

ataaagtttt	actgggacaa	agccaagcct	atcatttgca	aattgtctat	aaatattttc	37920
atgttacaga	atcacacagt	ttcaacagag	accatcttgt	ctacaaagct	gaaaatatct	37980
actatctggc	ccttgaagaa	agtttgccaa	accttagttt	atataataaa	agatcagcta	38040
tctcatagac	acctatctca	cacaacacat	tgtgggaaag	gaccttcttt	tttttttgag	38100
acggggtctt	gctctgttga	ccaggctgga	ctgtagtggc	atgatcatgg	ctcactgcag	38160
cctcaacctc	ccaggttcaa	gtaatgctcc	caccacagaa	tcccaaacag	ctgggagaga	38220
tgtgtgccac	tacgcctggc	taaggggcct	ttttaacaga	gaaagaaatc	cacatactac	38280
taagaaaaag	aaggggcatat	ttgatataata	tttatatttt	ttatatagat	atcataaaaa	38340
tcaagatgaa	ttatacagtt	atattttgca	atgtgtttga	cggtaaaagt	ttaatatcta	38400
taaaaattat	tttataaaat	atctttaata	tatttataga	tattataata	taaaatatct	38460
ataaaattat	tttataaaat	aaaaagttaa	gaagaaaaga	taggcaaaac	aaaaatcacg	38520
gcaatttaca	gaaaaccaag	tccaatgtgt	caacaaagat	aaaacagatt	tataaaactca	38580
ctaagtgtga	gagaattatt	agttaaagta	aaaatatctc	tctataccca	caatactact	38640
aaaaatcaga	gttataatgc	cctattgctg	gtggagatgt	aaggggagaa	gcattgctctc	38700
atatactgtt	agtgaataat	taactaata	catttttgaa	aagtaagctg	gcaatttttt	38760
ttttaactct	taccttttga	tgcaaaaact	catttttggg	tacctattcc	ataccttaaa	38820
aaaaatacat	atgcttactg	tagtactgtt	tataatggta	aaaactagaa	aaaaagaaaa	38880
cttgatagtg	aatactgaac	aaattacagt	gcattctacag	attaaacata	atgcagccat	38940
taaaaaagaa	taaattaggc	tgggtgcggt	ggctcatgcc	cgtaatccca	gcactttggg	39000
aggccaaagc	agtcggatca	cttgaggcca	ggagtctcag	accagcctgg	ccaacatggc	39060
aaaaccctgg	ctctacaaaa	aatacaaaaa	ttagtggggc	atgggtgggg	gcacctgtag	39120
tcccagctac	tcaggaggct	gaggcaggag	aatcacttga	gcctgggaga	cagagattgc	39180
agtgagccaa	gatcatgccca	cagcattcca	gtccagggtga	cagaacgaga	ctctgtctca	39240
acaaaaagaa	caaatttaaac	cctacaactc	atcaacaaaa	atacccaaac	ccaattcaaa	39300
aatgggcaaa	ggacttgaat	agacatttct	tcaaggatga	taaacaagca	catgaaaaaga	39360
tgcagagcac	tattcattag	tgattacatc	ccacatgcat	taggatggct	agtatgaaga	39420
acagaaaaata	ataaatattg	gtgaagatct	gaaaaacaga	aacctttgtg	cactgttggt	39480
gggaattgtaa	agtggtacag	ctactacgga	aaacagtatg	gccattcctc	aagaaaaata	39540
aaataaaatt	atcttatgat	aggaatatgc	atttctgggt	aaatacccca	aataactgaa	39600
aacagggtgt	acacccattt	caacatttac	atgtcaattc	aactgggcca	gaataccag	39660
atatttgttc	aaatattctt	ctggatgctt	ctatatatat	gttttttggc	tgagggttaac	39720
attttaaattg	gtgattctg	agtacagcag	attaccatcc	acaatgtagg	tgggcctcat	39780
ctactcagtt	gaaggcttta	cagaaaaaga	ctgacctccc	ttgagcaaga	aagaattcag	39840
gcaacagact	gcctttggac	tcaactgcaa	ctcttccttg	agtcaacagc	ccatcccatc	39900
accctggctt	ggtgagtcga	gggtctgatg	aggtaggctg	cagactcaag	gaagagctgc	39960
caaaaccagc	aaagccaatt	cattaaaata	aatctctctc	tacacaaaaca	cacacacaca	40020
ctaccaccac	caccatgatg	gttctgtttc	tctggagaat	gctaatacac	ccctgttcat	40080
ggcagcatta	ttcacaatag	ccaaaagggt	gaagcaactc	cagcagatga	atggagaagc	40140
aaaaatgtgg	atgtatatat	aatggaatat	tattaagcct	ttaaaaagtg	gaaattatat	40200
ctatctatat	ctatacacac	atactcacac	acacacacac	acatttatag	aagacagggt	40260
ttcaccatgt	tgtcaaggct	ggtctcgaac	tcctgggctc	aagcaaaccc	cctgcctcag	40320
cttcccaaag	tgctgagatt	acatgtgtga	gccaccacac	ccagccaaaa	aaaggacatt	40380
ctgacacata	atacaatata	gataaacaat	gaggacatca	tgatatgcga	aataagcctg	40440
tcacaaaaag	gcaattagtg	tatgattcct	cttgatagag	gtacctatgg	atgtcaaatc	40500
cataaaagtag	aatggggaaa	cagagagttg	tttaatgggt	atagagtttg	ttttgcaaga	40560
agaaaaagagt	tttgagaaat	gaatgtacaa	cagtgtgaac	ataatttaaca	ctactgaaaa	40620
tggttaagat	tataaatttt	atgttacatt	tattttacca	tgattaaaaa	ttaaaacaaa	40680
ataatattaa	ggaaaaatac	tataaataac	aacaacaaaa	aaaacacctc	aagcaactta	40740
cattcacctg	ggaaacagaa	tacatcctat	tctgctagag	atatatctgc	agttcaaaat	40800
ttattacaaa	tgatgttggt	tatctttttg	aaatgactga	aaaactaaat	taaaagcaat	40860
aatattcagt	ttactaacca	gtaagtcctt	ctttcatggt	tcctgacttt	tctgtaagat	40920
gttattgcaa	gatattctact	aaaatggaaa	acaactgaaa	aggcaaaaatt	ataatttctt	40980
atcaacatcg	ctaaaaccct	ggaggggaag	aatcctaaca	aacatggcca	taatttgcca	41040
catatttcta	ctgtcctcac	ttttcaaaat	ccagaaatca	acatttctgg	aaacaaaaaca	41100
gagtcataaa	tttggtcctt	tcttcagttt	agaagggtgc	aagttaatcc	ctgacatcct	41160
agttttccatt	ttcaaaaaatg	tactttttct	ctccccaaac	cggtatctag	attctttaat	41220
attttttagca	catagaagtt	aaatagattt	gcttaaccaa	aatagccagt	aaacctccca	41280
aaagaattaa	aatattaatg	gcgctttaat	gatacaaatg	aacaacttta	cattcaatcg	41340

tcaatgggaa	aggaagcaga	attctgagga	ttatgaaagt	aaacaaaacg	aagttcaaat	41400
tctactttat	tttacttttt	tgtaactaat	gaacaacttc	ttccaaagac	aagtaggaaa	41460
tacaaaaatt	agccaggcat	ggcacatgcc	tgtagtcctg	gttacttgga	aggctgaagt	41520
gggtggatcg	cttgagccgg	gaaggcagag	gctgtagtga	gctgagatca	catcacgcga	41580
ctcaagcctg	ggtgacagag	caagaccctc	tctggggaaa	aaaaaaaaaa	aaataggctg	41640
ggcgcagtg	ctcacacttg	taattccagc	actttgggag	gctgaggcag	gtggttcacc	41700
tgaggtcagg	agttctagac	cagcctgacc	aatatggtga	aaccctgtct	ctactaaaaa	41760
tacaaaaatt	agccaggcat	ggtggtgggc	aattgtaatc	ctagctactc	gggaggtcga	41820
ggcaggaaaa	tcgcctgaac	ccaagaggcg	gaggtttcag	tgagccgaga	ttgcactagt	41880
gcactccagc	ctgggcgaca	gagcaagact	tcatctcaaa	ataaataaat	aagtaagtaa	41940
ataaaaattaa	aaaatatata	aaaataaaac	aaagataagt	aggaaccatc	cttttttttt	42000
tttttttttt	ttttttttta	agatagggtc	tgtttctgat	gccagggtct	gagtgtagt	42060
gcatgatcat	ggctcactgc	aaccttgacc	tctcaaatac	aagtgactct	cctacctcag	42120
cctcccaagt	agctgggact	acaggtgctt	accaccccat	ccggctcat	taaaaaaat	42180
tttttgtaga	ggtgggggtc	cactatgttg	tatccaggct	ggtctcatt	taactttat	42240
agaaaaacag	catgtgttta	tcagcttctt	gtttttttta	aactaaaaat	aacactgttg	42300
ggttggttct	atgaagattc	tctaaattta	tttataacct	taagaataac	atgtagaaca	42360
aagtagatga	ctgaatgatc	tttggtgaat	aaatatgaat	ggatattcaa	ataatcaaaa	42420
atctcttaag	atctcccat	ctttacagga	tacagagaaa	actcgtaaat	atggcctgac	42480
ttttaccttt	gcagccttat	ccaaactctg	tggtcaagac	aaacagggtg	tccttatatc	42540
tacaacgtcc	ccctttgcct	acaaagctct	tctcatgact	ctttgcctat	cttaagttca	42600
cctatctgtc	aaatctctg	gaatgcaaca	tttctcaag	gtagccttct	ctcctcccaa	42660
actagaacaa	attcttctctg	gggcattagg	tttttattgc	actgtatgtc	tcttcttcac	42720
agcaatcaac	tggtcaatgt	tatatattga	ttcttagttg	atttggttct	ttccaccttt	42780
agactataac	cttctaagg	gtcacacata	atatcgatca	tcagttgtat	cccttggtga	42840
tagcacaggg	catggcaggc	aaatatgtgt	gtaaataaac	ttgttggaatg	aatcaatgag	42900
acacactttt	cttaccctaa	gtataatggc	aggataaac	ttatcaatct	attgcttctt	42960
gaaaaacaga	tatgatgtgc	tttaatttca	ttttacatct	caaataccaa	tgccctaagg	43020
attcacagtc	attttacaaa	tctttttgac	aaatgccttc	attaatcacc	acctgtttac	43080
aagtgtctaa	taacattttg	gttacattct	gtaacatttc	ctgcacttaa	tgtcatctct	43140
agaatactgg	ctaatatgaa	gcacctggac	ttcaggaaca	caaacctgaa	actaacacac	43200
caaactaaac	tggtatgtaa	atgacagaaa	tgacacattt	tggtctgcaa	catctctaga	43260
tggtcttttg	accaattcaa	cttttaccac	taaaaaatcg	tcacctgact	atagtcattt	43320
tgagctcatg	ataaatgaat	tacagatgaa	aaataaatag	ttgatgaca	atctttacaa	43380
aagtttatct	tcaaagaata	ccaccagtca	caggtattct	aggctcctat	caactttatt	43440
ggtcagggca	gaattcactt	ttcatgataa	ttatgttctg	aaaattctac	aaacttaatt	43500
attacaaaaca	aaagtcatag	tttgctcata	aatcaggcct	aggctctggat	tctagtctct	43560
ccattttttca	ttgtttcact	gaggcaagtg	acttaaaatt	ccctagcctc	agtttctctc	43620
catgtaaaat	cagataatga	ttcctattcc	taagatgggt	ttgaggcttc	aacaagataa	43680
gatgggctc	actcaagcat	gctcagtact	ctgtctctct	ctctccggtt	atgcagaaat	43740
tctatttagga	ttctgcaaag	taaaaataat	atttcagtaa	aaattatgcc	ctttattaat	43800
gaatctagat	tttcagattt	tccttaaaatt	tacttagtaa	cttaagggtc	caaatattat	43860
agagatttgt	atctagtatt	ttaaagaaat	gaaagggtgt	aatcaaaaatg	ctgcacaaat	43920
aaaatgtaca	tttaacaaac	agaatatcac	aaccatacaa	actaatcaga	tataaagaag	43980
tcagcaacag	aaatctgatg	ttgcctttag	atcacacaat	taggcaaaaca	aaaatagagt	44040
tccatcctcc	tttggtcaag	gccatgggtg	aagactgaat	accaaatagg	gaaaataggaa	44100
aagccaggaa	atggcaaat	agcaaaaact	ggactcctta	attttttat	tcattttcat	44160
atctcacttc	taaaaacttta	attaaattca	aataaaaacc	aaaatggaac	tgagataaag	44220
ccaaaaggaa	agttatgtag	gtcaaatgag	aacctatatt	gtccttaggc	tctttgttgc	44280
tttctgttta	aggaaaaact	gcccaagtgc	cttgacacat	taaagatcaa	gcaggagggt	44340
ctgccgagag	tcctccatctg	gcagccaggt	tttgtcaagc	aaattttgag	aattctctac	44400
cctcccactt	tctatctaatt	tatagcactt	tataaaaacc	attctctctc	tgtctctgtc	44460
tctctctctc	tctctctctc	acacacacac	acacacacac	acacacacac	acacacaccc	44520
tttctctctc	tctctctctg	aaacttatct	gtattataat	aacacaacac	taggtatgga	44580
ttaatctgac	aattttcccc	taaaacagaa	taaattcaaa	aaggaaaacc	tttctctctg	44640
acacatgcac	tatattctga	caataataat	tcctaaaatta	agtataatac	attttcccta	44700
caggagttaa	aagaagttac	agtaaagaat	ctcttgata	aatatataatg	ccagaacttg	44760
acccaaataa	gtgctgagag	gtataaatct	caaaacagtt	tccggactct	ttgtgaaatg	44820

```

tcttcagagt ctgcgatata ttttcttcaa ctaaattata caagtaagat attttgctgg 44880
gctgtgggaa tgccttacgg catgttactg tggagctcat ggtaaaatag aaagaatata 44940
aataattaaa ataaaattga caaatgataa atgatttaat aaattagaaa ttcaaatgcc 45000
gggcactttt ctagaacctg gacacaaaagc atgaacctaa caataacccc gccttcctga 45060
aaaatatgga ctatttgaaa attatactcg caacactaaa taaatattct tcattctctc 45120
agtatatgga gatgtttact ttcaattaga caatttgctt tcctctctga acacatagtt 45180
atgtgatggc tctataaaaag attttaaaat aactatagaa ggaactattg gtaaagactg 45240
tgggatacta aaaatggcta caaagaaaag tatgacaaaa cctctgagtt tgaatggaag 45300
tcctactaga ttagagtcta agcctgtgac attatgcttc tgggtcttgt tcttaaatgc 45360
ttttctcatt aatagtatgt aacttacttc ctggaatgcc attcattaaa aaaatattta 45420
atatttgcta aatgtcaata tttatgccag cactttttaa gtacagaaac atggagtttc 45480
tttacctcat gcaaatatgc tgtgagaaag acttaagagc ctattgccta ctttgtggta 45540
caacactgaa gactcaccat ccaaaacaaa cagacttagt aaattcttgt gatttgcagt 45600
agttctgttc tataaggtta ccacaaacac tgaatcatc gctcctgggg gaatacaagg 45660
ttatgttttc gtgagccctc ggtcacaaca tgttcattaa ctgatcaata cataaccttg 45720
ttctatgtgt gtttctgttt aaaaagagca cttcagtgtc acatttggag tctgttttaa 45780
acagcaaaat cactaataaa aagcacaaaa atgtaaaagc atggcactac atacactgtg 45840
acaagaaggc ttgtttatag tatgacagct gagacaagaa ggtagagcct cgctttgatc 45900
aacctctgct gggaaatgag catcaggtga atcaattttt caccactctg aatgaccgta 45960
aaagtgtctc aagtactgac tttgggttta cacataaatt ttagtaagca tgtgaatctg 46020
ccaatatgaa atctacaaat aatgagtacc aaatgcatat gagtcaaata tttcagtgcg 46080
gtatctgact tgattgccac tgaaagacac agtttggaat acccctaata aataccgttt 46140
agttactatg cagacaaaaga gttctacact agagtgtctc aattaagatg tctgaggctt 46200
tcataaatgg atgtttttta aaatgttatc tcctacctga tatattctaa aggggatata 46260
acgaaatcca ttttcttctg caggatattc catgagtttc cgattgatgg cccaaaactg 46320
gtcaaatctg tctgtaatga 46340

```

<210> 67  
 <211> 773  
 <212> ADN  
 <213> Homo sapiens

```

<400> 67
actgagagac aggactagct ggatttccta ggctgactaa gaatccctaa gcctagctgg 60
gaaggtgacc acatccacct ttaaacacgg ggcttgcaac ttagctcaca cctgaccaag 120
gaaggtgacc acaccctcct ttaaacacag agcttgtaac tcagctcaca cccgaccaat 180
caggtagtaa agagagctca ctaaaatacc aattaggcta aaaacaggag gtaaaagaaat 240
aatcaaatca tctatcgctt gagagcacag ggggaggagc aatgatcggg atataaacc 300
aggcatttga gccagatcag gtaaccctct ttgggtcccc tcacactgta tgggagctct 360
gttttcactc tattaatatc tgcaactgca cactcttctg gtccatgttt gttccggctc 420
aagctgagct tttgctcgcc gtccaccact gctgaatgcc gccattgcag acctgccctt 480
gacttccacc cctccggatc cggcagagtg tccgctgcac tcctgatcca gcgaggcacc 540
cattgccact cccgatcagg ctaaaaggctt gccattgttc ctgcacagct aagtgcctgg 600
gttcactcta atcaggctga acactggctg ctgggttcca cggttctctt ccatgactca 660
cagcttctaa tagagctata acactacca catggcccaa ggttccattc gttggaatcc 720
atgaggccaa gaaccccagg tcagagaata aaaggcccgcc cccatcttgg gag 773

```

<210> 68  
 <211> 10  
 <212> PRT  
 <213> Homo sapiens

```

<400> 68
Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val
1           5           10

```



<210> 69  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 69  
Leu Leu Phe Gly Pro Cys Ile Phe Asn Leu  
1 5 10

<210> 70  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 70  
Cys Leu Pro Leu Asn Phe Arg Pro Tyr Val  
1 5 10

<210> 71  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 71  
Gly Leu Leu Ser Gln Trp Met Pro Trp Ile  
1 5 10

<210> 72  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 72  
Cys Leu Pro Ser Gly Ile Phe Phe Val  
1 5

<210> 73  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 73  
Trp Met Pro Trp Ile Leu Pro Phe Leu  
1 5

74

<210> 74  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 74  
Ile Arg Trp Val Thr Pro Pro Thr Gln Ile  
1 5 10

<210> 75  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 75  
Leu Arg Asn Thr Gly Pro Trp Gly Leu Leu  
1 5 10

<210> 76  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 76  
Leu Arg Thr His Thr Arg Leu Val Ser Leu  
1 5 10

<210> 77  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 77  
Lys Arg Val Pro Ile Leu Pro Phe Val Ile  
1 5 10

<210> 78  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 78  
Cys Arg Cys Met Thr Ser Ser Ser Pro Tyr  
1 5 10

75

<210> 79  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 79  
Thr Arg Val His Gly Thr Ser Ser Pro Tyr  
1 5 10

<210> 80  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 80  
Ala Arg Glu Lys His Val Lys Glu Val Ile  
1 5 10

<210> 81  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 81  
Ser Arg Ile Glu Ala Val Lys Leu Gln Met  
1 5 10

<210> 82  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 82  
Ser Gln Trp Met Pro Trp Ile Leu Pro Phe  
1 5 10

<210> 83  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 83  
Cys Tyr Tyr Val Asn Gln Ser Gly Ile  
1 5

<210> 84  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 84  
Phe Tyr Tyr Lys Leu Ser Gln Glu Leu  
1 5

<210> 85  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 85  
Thr Tyr Thr Thr Asn Ser Gln Cys Ile  
1 5

<210> 86  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 86  
Ser Phe Leu Val Pro Pro Met Thr Ile  
1 5

<210> 87  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 87  
Tyr Tyr Val Asn Gln Ser Gly Ile Val  
1 5

<210> 88  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 88  
Leu Phe Asn Thr Thr Leu Thr Gly Leu  
1 5

77

<210> 89  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 89  
Leu Phe Gly Pro Cys Ile Phe Asn Leu  
1 5

<210> 90  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 90  
Arg Trp Val Thr Pro Pro Thr Gln Ile  
1 5

<210> 91  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 91  
Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile  
1 5 10

<210> 92  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 92  
Leu Pro Tyr His Ile Phe Leu Phe Thr Val  
1 5 10

<210> 93  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 93  
Gly Ala Leu Gly Thr Gly Ile Gly Gly Ile  
1 5 10

78

<210> 94  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 94  
Leu Pro Phe Val Ile Gly Ala Gly Val Leu  
1 5 10

<210> 95  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 95  
Arg Arg Pro Leu Asp Arg Pro Ala Ser  
1 5

<210> 96  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 96  
Phe Arg Pro Tyr Val Ser Ile Pro Val  
1 5

<210> 97  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 97  
Arg Arg Ala Leu Asp Leu Leu Thr Ala  
1 5

<210> 98  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 98  
Trp Arg Met Gln Arg Pro Gly Asn Ile  
1 5

79

<210> 99  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 99  
Asp Arg Ile Gln Arg Arg Ala Glu Glu Leu  
1 5 10

<210> 100  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 100  
Leu Arg Thr His Thr Arg Leu Val Ser Leu  
1 5 10

<210> 101  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 101  
Glu Arg Val Ala Asp Ser Leu Val Thr Leu  
1 5 10

<210> 102  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 102  
Leu Phe Gly Pro Cys Ile Phe Asn Leu Leu  
1 5 10

<210> 103  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 103  
Gln Phe Tyr Tyr Lys Leu Ser Gln Glu Leu  
1 5 10

80

<210> 104  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 104  
Gln Trp Met Pro Trp Ile Leu Pro Phe Leu  
1 5 10

<210> 105  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 105  
Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val  
1 5 10

<210> 106  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 106  
Asn Phe Val Ser Ser Arg Ile Glu Ala Val  
1 5 10

<210> 107  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 107  
Gly Pro Leu Val Ser Asn Leu Glu Ile  
1 5

<210> 108  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 108  
Leu Pro Leu Asn Phe Arg Pro Tyr Val  
1 5



81

<210> 109  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 109  
Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile  
1 5 10

<210> 110  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 110  
Glu Pro Lys Met Gln Ser Lys Thr Lys Ile  
1 5 10

<210> 111  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 111  
Leu Pro Tyr His Ile Phe Leu Phe Thr Val  
1 5 10

<210> 112  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 112  
Arg Glu Lys His Val Lys Glu Val Ile  
1 5

<210> 113  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 113  
Lys Pro Arg Asn Lys Arg Val Pro Ile Leu  
1 5 10

82

<210> 114  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 114  
Val Val Leu Gln Asn Arg Arg Ala Leu  
1 5

<210> 115  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 115  
Ala Val Val Leu Gln Asn Arg Arg Ala Leu  
1 5 10

<210> 116  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 116  
Leu Pro Phe Val Ile Gly Ala Gly Val  
1 5

<210> 117  
<211> 9  
<212> PRT  
<213> Homo sapiens

<400> 117  
Asp Leu Tyr Ser Tyr Val Ile Ser Lys  
1 5

<210> 118  
<211> 10  
<212> PRT  
<213> Homo sapiens

<400> 118  
Thr Glu Gln Asp Leu Tyr Ser Tyr Val Ile  
1 5 10

<210> 119  
<211> 2615  
<212> ADN  
<213> Homo sapiens

<400> 119  
gaattccggg aagccagacg gttaacacag acaaagtgtc gccgtgacac tcggccctcc 60  
agtgttgagg agaggcaaga gcagcgaccg cgcacctgtc cgcccgagc tgggacgcgc 120  
gcccgggagg cgggacgaag cgaggaggga ccgcccagc tgccccaag tgtaactcca 180  
gcactgtgag gtttcaggga ttggcagagg ggaccaaggg gacatgaaaa tggacatgga 240  
ggatgaggat atgactctgt ggacagaggc tgagtttgaa gagaagtgtg catacattgt 300  
gaacgaccac ccctgggatt ctggtgctga tggcggtagt tcggttcagg cggaggcatc 360  
cttaccagg aatctgcttt tcaagtatgc caccaacagt gaagagggtt ttggagtgt 420  
gagtaaaagaa tacataccaa agggcacacg ttttgacccc ctaatagggtg aaatctacac 480  
caatgacaca gttcctaaga acgccaacag gaaatatatt tgaggatct attccaggg 540  
ggagcttcac cacttcattg acggctttta tgaagagaaa agcaactgga tgcgctatgt 600  
gaatccagca cactctcccc gggagcaaaa cctggctgag tgtcagaacg ggatgaacat 660  
ctacttctac accattaagc ccatccctgc caaccaggaa cttcttctgt ggtattgtcg 720  
ggactttgca gaaaggcttc actaccctta tcccggagag ctgacaatga tgaatctcac 780  
acaaacacag agcagtctaa agcaaccgag cactgagaaa aatgaactct gcccaaagaa 840  
tgtcccaaag agagagtaca gcgtgaaaag aatcctaaaa ttggactcca acccctccaa 900  
aggaaaggac ctctaccgtt ctaacatttc acccctcaca tcagaaaagg acctcgatga 960  
ctttagaaga cgtgggagcc ccgaaatgcc cttctaccct cgggtcggtt accccatccg 1020  
ggccctctct ccagaagact ttttgaaagc ttccctggcc tacgggatcg agagaccac 1080  
gtacatcact cgctccccc tccatcctc caccactcca agccctctg caagaagcag 1140  
ccccgaccaa agcctcaaga gctccagccc tcacagcagc cctgggaata cgggtgtccc 1200  
tgtgggcccc ggtctctcaag agcacggga ctctacgct tacttgaacg cgtcctacg 1260  
cacggaaggt ttgggtcct accctggcta cgccccctg cccacctcc gccagcttt 1320  
catccctctg tacaacgctc actaccctaa gttctcttg cccctctac gcatgaattg 1380  
taatggcctg agcgtgtga gcagcatgaa tggcatcaac aactttggcc tcttcccag 1440  
gctgtgccct gtctacagca atctcctcgg tgggggcagc ctgccccacc ccatgctcaa 1500  
ccccacttct cccccagct cgctgccctc agatggagcc cggaggttgc tccagccgga 1560  
gcatccagc gaggtgcttg tcccggcgcc ccacagtgc ttctcctta cggggccgc 1620  
cgccagcatg aaggacaagg cctgtagccc cacaagcggg tctccacag cgggaacagc 1680  
cgccacggca gaacatgttg tgcagcccaa agctacctca gcagcgatg cagccccag 1740  
cagcgacgaa cagcatgaatc tcattaaaaa caaaagaaac atgaccggt acaagacct 1800  
tccctaccct ctgaagaagc agaacggcaa gatcaagtac gaatgcaacg tttgcgcaa 1860  
gactttcggc cagctctcca atctgaaggt ccacctgaga gtgcacagt gagaacggcc 1920  
tttcaaattg cagacttgca acaagggtt tactcagct gccacctgc agaaacacta 1980  
cctggtacac acgggagaaa agccacatga atgccaggtc tgccacaaga gatttagcag 2040  
caccagcaat ctcaagacc acctgcgact ccattctgga gagaaccat accaatgcaa 2100  
ggtgtgccct gccaaagtta ccagtttgt gcacctgaaa ctgcacaagc gtctgcacac 2160  
ccgggagcgg ccccaacagt gctcccagt ccacaagaac tacatccatc tctgtagct 2220  
caaggttcac ctgaaaaggga actgcgctgc ggccccggcg cctgggtgc cctgggaaga 2280  
tctgaccgca atcaatgaag aaatcgagaa gtttgacatc agtgacaatg ctgaccggt 2340  
cgaggacgtg gaggatgaca tcagtgtgat ctctgtagt gagaaggaaa ttctggcct 2400  
ggtcagaaaa gagaaagaag aaactggcct gaaagtgtct ttgcaaagaa acatggggaa 2460  
tggactctc tcctcagggt gcagccttta tgagtcatca gatctacccc tcatgaagt 2520  
gcctccagc aaccactac ctctggtacc tgtaaaggct aaacaagaaa cagttgaacc 2580  
aatggatcct taagattttc agaaaacact tattt 2615

<210> 120  
<211> 29  
<212> PRT  
<213> Homo sapiens

84

&lt;400&gt; 120

Leu Gln Asn Arg Arg Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly  
1 5 10 15

Thr Cys Leu Phe Leu Gly Glu Glu Cys Cys Tyr Tyr Val  
20 25

&lt;210&gt; 121

&lt;211&gt; 21

&lt;212&gt; ADN

&lt;213&gt; Homo sapiens

&lt;400&gt; 121

cttcaaaca caaccaggag g

21

&lt;210&gt; 122

&lt;211&gt; 20

&lt;212&gt; ADN

&lt;213&gt; Homo sapiens

&lt;400&gt; 122

ttggggaggt tggccgacga

20

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 99/01513

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC 6 C12N15/48 C12Q1/70 C07K14/15 A61K31/70		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) IPC 6 C12N C12Q C07K A61K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 23755 A (BIO MERIEUX) 4 June 1998 (1998-06-04) Comparez nucléotides 1-1462 de SEQ ID NO:117 avec nucléotides 928-2390 de SEQ ID NO:1 de la présente demande; comparez SEQ ID NO:118 avec SEQ ID NOs:22 et 120 de la présente demande ---	1,3-12, 14-36
X	Database GenBank. Séquence HSAC 000064 Clone humain BAC RG083M05 de 7q21-7q22, séquence complet. 17 novembre 1996. XP002118730 Comparez nucléotides 28000-38500 d' AC00064 avec SEQ ID NO:3 --- -/--	1-4,13
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "Z" document member of the same patent family		
Date of the actual completion of the international search 18 October 1999		Date of mailing of the international search report 11.11.99
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer Cupido, M

# INTERNATIONAL SEARCH REPORT

Internal Application No

PCT/FR 99/01513

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>ALLIEL PM ET AL: "Séquences rétrovirales endogènes anlogues à celle du nouveau rétrovirus MSRV associé à la sclérose en plaques ( première partie)"            COMPTES RENDUS DES SEANCES DE L'ACADEMIE DES SCIENCES SERIE III: SCIENCES DE LA VIE.,            vol. 321, no. 6, June 1998 (1998-06),            pages 495-499, XP002101380            MONTREUIL FR            figures 2,3</p> <p>---</p>	1,3-12, 14-36
X	<p>Database GenBank Séquence AC X93499            mRNA de H. sapiens pour la protéine rab7            10 février 1997            XP002119234            &amp; VITELLI R ET AL: "Molecular cloning and expression analysis of the human rab7 GTP-ase complementary deoxyribonucleic acid"            BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS,            vol. 229, no. 3, 1996, pages 887-890,            ORLANDO, FL US</p> <p>---</p>	1-4
X	<p>FR 2 737 500 A (BIO MERIEUX)            7 February 1997 (1997-02-07)            cited in the application            the whole document</p> <p>---</p>	1,3-12, 14-36
P,X	<p>WO 99 02666 A (BIO MERIEUX)            21 January 1999 (1999-01-21)            Comparez SEQ ID NOs 130, 117, 114 et 120            avec SEQ ID NOs: 1-3 de cette demande et            SEQ ID NO:118 avec SEQ ID NOs:22 et 120 de            cette demande</p> <p>---</p>	1,3-12, 14-36
P,X	<p>WO 99 02696 A (BIO MERIEUX ;BESEME            FREDERIC (FR); BLOND JEAN LUC (FR); BOUTON            OLI) 21 January 1999 (1999-01-21)            Comparez SEQ ID NOs: 4, 5, 7, 9 et 11 avec            SEQ ID NOs: 1-3 de cette demande.</p> <p>---</p>	1,3-12, 14-36
P,X	<p>ALLIEL PM ET AL: "Rétrovirus endogènes et            sclérose en plaques. II. HERV-7q"            COMPTES RENDUS DES SEANCES DE L'ACADEMIE            DES SCIENCES SERIE III: SCIENCES DE LA            VIE.,            vol. 321, no. 10, October 1998 (1998-10),            pages 857-863, XP002101381            MONTREUIL FR            the whole document</p> <p>---</p>	1,3-12, 14-36
	<p>---</p> <p>-/--</p>	

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/FR 99/01513

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 99 26972 A (GENETICS INSTITUTE, INC.) 3 June 1999 (1999-06-03) comparez SEQ ID NO:4 avec a.a. 131-668 de séquence 22 de cette demande ---	1-4
A	MITANI M ET AL: "Suppressive effect on polyclonal B-cell activation of a synthetic peptide homologous to a transmembrane component of oncogenic retrovirus" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 84, no. 1, January 1987 (1987-01), pages 237-240, XP002118729 WASHINGTON US cited in the application the whole document -----	24,25

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/FR 99/01513

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☒ Claims Nos.: 8  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

See supplementary sheet INFORMATION FOLLOW-UP PCT/ISA/210

3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See supplementary sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.



The International Searching Authority found several (groups of) inventions in the international application, namely:

1. Claims: 1, 9, 21-23, 26 (wholly), 28, 10-20 and 27-37 (partly)

Nucleic acid fragments derived from HERV-7q env, diagnostic reagents, diagnostic applications and kits, peptides, pharmaceutical compositions and applications, antibodies and corresponding transgenic animals.

2. Claims: 2-7, 10-20, 27-37 (all partly)

Nucleic acid fragments derived from HERV-7q gag, diagnostic reagents, diagnostic applications and kits, peptides, pharmaceutical compositions and applications, antibodies and corresponding transgenic animals.

3. Claims: 2-7, 10-20, 27-37 (all partly)

Human nucleic acid fragments similar to HERV-7q gag, diagnostic reagents, diagnostic applications and kits, peptides, pharmaceutical compositions and applications, antibodies and corresponding transgenic animals.

4. Claims: 24, 25

Compositions comprising a CKS-type motif, inasmuch as said compositions do not contain a sequence as per the first invention.

Continuation of Box I.2

Claim No: 8

Claim 8 concerns a very wide variety of compounds. A support basis as defined in PCT Article 6 and a description as defined in PCT Article 5 can however be found for only a very limited number of the claimed compounds. In the present case, the claims are so lacking in support basis and the disclosure of the invention in the description is so limited that it is impossible to carry out any significant search concerning the whole claimed spectrum.

The applicant's attention is drawn to the fact that claims concerning inventions in respect of which no search report has been established need not be the subject of a preliminary examination report (PCT Rule 66.1 (e)). The applicant is warned that the guideline adopted by the EPO acting in its capacity as International Preliminary Examining Authority is not to proceed with a preliminary examination of a subject matter unless a search has been carried out thereon. This position will remain unchanged, notwithstanding that the claims have or have not been modified, either after receiving the search report, or during any procedure under Chapter II.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/FR 99/01513

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9823755 A	04-06-1998	EP 0942987 A	22-09-1999
FR 2737500 A	07-02-1997	AU 6823296 A	05-03-1997
		BG 101355 A	30-12-1997
		BR 9606566 A	30-12-1997
		CA 2201282 A	20-02-1997
		CZ 9701357 A	17-06-1998
		EP 0789077 A	13-08-1997
		WO 9706260 A	20-02-1997
		HU 9900425 A	28-05-1999
		JP 11502416 T	02-03-1999
		NO 971493 A	03-06-1997
		NZ 316080 A	29-04-1999
		PL 319512 A	18-08-1997
		SK 56797 A	09-09-1998
WO 9902666 A	21-01-1999	FR 2765588 A	08-01-1999
		AU 8545098 A	08-02-1999
WO 9902696 A	21-01-1999	AU 8447098 A	08-02-1999
WO 9926972 A	03-06-1999	AU 1417899 A	15-06-1999

# RAPPORT DE RECHERCHE INTERNATIONALE

Dem: Internationale No

PCT/FR 99/01513

<b>A. CLASSEMENT DE L'OBJET DE LA DEMANDE</b> CIB 6 C12N15/48 C12Q1/70 C07K14/15 A61K31/70		
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB		
<b>B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE</b> Documentation minimale consultée (système de classification suivi des symboles de classement) CIB 6 C12N C12Q C07K A61K		
Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche		
Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)		
<b>C. DOCUMENTS CONSIDERES COMME PERTINENTS</b>		
Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	WO 98 23755 A (BIO MERIEUX) 4 juin 1998 (1998-06-04) Comparez nucléotides 1-1462 de SEQ ID NO:117 avec nucléotides 928-2390 de SEQ ID NO:1 de la présente demande; comparez SEQ ID NO:118 avec SEQ ID NOs:22 et 120 de la présente demande ---	1,3-12, 14-36
X	Database GenBank. Séquence HSAC 000064 Clone humain BAC RG083M05 de 7q21-7q22, séquence complet. 17 novembre 1996. XP002118730 Comparez nucléotides 28000-38500 d' AC00064 avec SEQ ID NO:3 --- -/--	1-4,13
<div style="display: flex; justify-content: space-between;"> <div> <input checked="" type="checkbox"/> Voir la suite du cadre C pour la fin de la liste des documents         </div> <div> <input checked="" type="checkbox"/> Les documents de familles de brevets sont indiqués en annexe         </div> </div>		
* Catégories spéciales de documents cités:		
<div style="display: flex;"> <div style="flex: 1;"> <p>"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent</p> <p>"E" document antérieur, mais publié à la date de dépôt international ou après cette date</p> <p>"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)</p> <p>"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens</p> <p>"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée</p> </div> <div style="flex: 1;"> <p>"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention</p> <p>"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément</p> <p>"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier</p> <p>"A" document qui fait partie de la même famille de brevets</p> </div> </div>		
Date à laquelle la recherche internationale a été effectivement achevée 18 octobre 1999		Date d'expédition du présent rapport de recherche internationale 11 11 99
Nom et adresse postale de l'administration chargée de la recherche internationale Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Fonctionnaire autorisé Cupido, M

# RAPPORT DE RECHERCHE INTERNATIONALE

Dem. Internationale No

PCT/FR 99/01513

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	<p>ALLIEL PM ET AL: "Séquences rétrovirales endogènes anlogues à celle du nouveau rétrovirus MSRV associé à la sclérose en plaques ( première partie)"</p> <p>COMPTES RENDUS DES SEANCES DE L'ACADEMIE DES SCIENCES SERIE III: SCIENCES DE LA VIE.,</p> <p>vol. 321, no. 6, juin 1998 (1998-06),</p> <p>pages 495-499, XP002101380</p> <p>MONTREUIL FR</p> <p>figures 2,3</p>	1,3-12, 14-36
X	<p>Database GenBank Séquence AC X93499</p> <p>mRNA de H. sapiens pour la protéine rab7</p> <p>10 février 1997</p> <p>XP002119234</p> <p>&amp; VITELLI R ET AL: "Molecular cloning and expression analysis of the human rab7 GTP-ase complementary deoxyribonucleic acid"</p> <p>BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS,</p> <p>vol. 229, no. 3, 1996, pages 887-890,</p> <p>ORLANDO, FL US</p>	1-4
X	<p>FR 2 737 500 A (BIO MERIEUX)</p> <p>7 février 1997 (1997-02-07)</p> <p>cité dans la demande</p> <p>le document en entier</p>	1,3-12, 14-36
P,X	<p>WO 99 02666 A (BIO MERIEUX)</p> <p>21 janvier 1999 (1999-01-21)</p> <p>Comparez SEQ ID NOs 130, 117, 114 et 120</p> <p>avec SEQ ID NOs: 1-3 de cette demande et</p> <p>SEQ ID NO:118 avec SEQ ID NOs:22 et 120 de cette demande</p>	1,3-12, 14-36
P,X	<p>WO 99 02696 A (BIO MERIEUX ;BESEME FREDERIC (FR); BLOND JEAN LUC (FR); BOUTON OLI) 21 janvier 1999 (1999-01-21)</p> <p>Comparez SEQ ID NOs: 4, 5, 7, 9 et 11 avec</p> <p>SEQ ID NOs: 1-3 de cette demande.</p>	1,3-12, 14-36
P,X	<p>ALLIEL PM ET AL: "Rétrovirus endogènes et sclérose en plaques. II. HERV-7q"</p> <p>COMPTES RENDUS DES SEANCES DE L'ACADEMIE DES SCIENCES SERIE III: SCIENCES DE LA VIE.,</p> <p>vol. 321, no. 10, octobre 1998 (1998-10),</p> <p>pages 857-863, XP002101381</p> <p>MONTREUIL FR</p> <p>le document en entier</p>	1,3-12, 14-36

-/--

# RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No  
PCT/FR 99/01513

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
P,X	WO 99 26972 A (GENETICS INSTITUTE, INC.) 3 juin 1999 (1999-06-03) comparez SEQ ID NO:4 avec a.a. 131-668 de séquence 22 de cette demande ---	1-4
A	MITANI M ET AL: "Suppressive effect on polyclonal B-cell activation of a synthetic peptide homologous to a transmembrane component of oncogenic retrovirus" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 84, no. 1, janvier 1987 (1987-01), pages 237-240, XP002118729 WASHINGTON US cité dans la demande le document en entier -----	24,25

# RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°  
PCT/FR 99/01513

## Cadre I Observations - lorsqu'il a été estimé que certaines revendications ne pouvaient pas faire l'objet d'une recherche (suite du point 1 de la première feuille)

Conformément à l'article 17.2)a), certaines revendications n'ont pas fait l'objet d'une recherche pour les motifs suivants:

1. ☐ Les revendications n°  
se rapportent à un objet à l'égard duquel l'administration n'est pas tenue de procéder à la recherche, à savoir:
2. ☒ Les revendications n° 8  
se rapportent à des parties de la demande internationale qui ne remplissent pas suffisamment les conditions prescrites pour qu'une recherche significative puisse être effectuée, en particulier:  
voir feuille supplémentaire SUITE DES RENSEIGNEMENTS PCT/ISA/210
3. ☐ Les revendications n°  
sont des revendications dépendantes et ne sont pas rédigées conformément aux dispositions de la deuxième et de la troisième phrases de la règle 6.4.a).

## Cadre II Observations - lorsqu'il y a absence d'unité de l'invention (suite du point 2 de la première feuille)

L'administration chargée de la recherche internationale a trouvé plusieurs inventions dans la demande internationale, à savoir:

voir feuille supplémentaire

1. ☐ Comme toutes les taxes additionnelles ont été payées dans les délais par le déposant, le présent rapport de recherche internationale porte sur toutes les revendications pouvant faire l'objet d'une recherche.
2. ☒ Comme toutes les recherches portant sur les revendications qui s'y prêtaient ont pu être effectuées sans effort particulier justifiant une taxe additionnelle, l'administration n'a sollicité le paiement d'aucune taxe de cette nature.
3. ☐ Comme une partie seulement des taxes additionnelles demandées a été payée dans les délais par le déposant, le présent rapport de recherche internationale ne porte que sur les revendications pour lesquelles les taxes ont été payées, à savoir les revendications n°
4. ☐ Aucune taxe additionnelle demandée n'a été payée dans les délais par le déposant. En conséquence, le présent rapport de recherche internationale ne porte que sur l'invention mentionnée en premier lieu dans les revendications; elle est couverte par les revendications n°

Remarque quant à la réserve

- ☐ Les taxes additionnelles étaient accompagnées d'une réserve de la part du déposant.
- ☐ Le paiement des taxes additionnelles n'était assorti d'aucune réserve.

## SUIITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA/ 210

Suite du cadre I.2

Revendications nos.: 8

Le revendication 8 présente à trait à une très grande variété de composés. Un fondement au sens de L'Article 6 PCT et un exposé au sens de l'Article 5 PCT ne peut cependant être trouvé que pour un nombre très restreint de ces composés revendiqués. Dans le cas présent, les revendications manquent à un tel point de fondement et l'exposé de l'invention dans la description est si limité q'une recherche significative couvrant tout le spectre revendiqué est impossible.

L'attention du déposant est attirée sur le fait que les revendications, ou des parties de revendications, ayant trait aux inventions pour lesquelles aucun rapport de recherche n'a été établi ne peuvent faire obligatoirement l'objet d'un rapport préliminaire d'examen (Règle 66.1(e) PCT). Le déposant est averti que la ligne de conduite adoptée par l'OEB agissant en qualité d'administration chargée de l'examen préliminaire international est, normalement, de ne pas procéder à un examen préliminaire sur un sujet n'ayant pas fait l'objet d'une recherche. Cette attitude restera inchangée, indépendamment du fait que les revendications aient ou n'aient pas été modifiées, soit après la réception du rapport de recherche, soit pendant une quelconque procédure sous le Chapitre II.



SUITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA/ 210

L'administration chargée de la recherche internationale a trouvé plusieurs (groupes d') inventions dans la demande internationale, à savoir:

1. revendications: 1, 9, 21-23, 26 (complet), 28, 10-20 et 27-37 (partiellement)

Fragments d'acide nucléique dérivé du HERV-7q env, réactifs de diagnostic, applications et kits diagnostiques, peptides, compositions et applications pharmaceutiques, anticorps et animaux transgéniques correspondants.

2. revendications: 2-7, 10-20, 27-37( tous partiellement)

Fragments d'acide nucléique dérivé du HERV-7q gag, réactifs de diagnostic, applications et kits diagnostiques, peptides, compositions et applications pharmaceutiques, anticorps et animaux transgéniques correspondants.

3. revendications: 2-7, 10-20, 27-37( tous partiellement)

Fragments d'acide nucléique humaines similaires à HERV-7q (SEQ ID NOs: 4-21 et 61), réactifs de diagnostic, applications et kits diagnostiques, peptides, compositions et applications pharmaceutiques, anticorps et animaux transgéniques correspondants.

4. revendications: 24, 25

Compositions comprenant un motif de type CKS, dans la mesure où ces compositions ne contiennent pas une séquence selon la première invention.

# RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale No

PCT/FR 99/01513

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9823755 A	04-06-1998	EP 0942987 A	22-09-1999
FR 2737500 A	07-02-1997	AU 6823296 A	05-03-1997
		BG 101355 A	30-12-1997
		BR 9606566 A	30-12-1997
		CA 2201282 A	20-02-1997
		CZ 9701357 A	17-06-1998
		EP 0789077 A	13-08-1997
		WO 9706260 A	20-02-1997
		HU 9900425 A	28-05-1999
		JP 11502416 T	02-03-1999
		NO 971493 A	03-06-1997
		NZ 316080 A	29-04-1999
		PL 319512 A	18-08-1997
		SK 56797 A	09-09-1998
WO 9902666 A	21-01-1999	FR 2765588 A	08-01-1999
		AU 8545098 A	08-02-1999
WO 9902696 A	21-01-1999	AU 8447098 A	08-02-1999
WO 9926972 A	03-06-1999	AU 1417899 A	15-06-1999